

Contract Title:

High-frequency electromagnetic propagation/scattering codes

Contract #: **F49620-03-C-0031**

STTR Phase II – Final Report

Period of Performance: **April 01, 03 – December 31, 05**

Prepared by:

MONOPOLE RESEARCH

739 Calle Sequoia, Thousand Oaks, CA 91360

tel: (805) 375-0318 fax: (805) 499-9878

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188	
Public Reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimates or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188,) Washington, DC 20503.				
1. AGENCY USE ONLY (Leave Blank)		2. REPORT DATE December 31, 2005		3. REPORT TYPE AND DATES COVERED Final Report, April 01, 03 – Dec 31, 05
4. TITLE AND SUBTITLE High Frequency Electromagnetic Propagation / Scattering Codes			5. FUNDING NUMBERS Contract F49620-03-C-0031	
6. AUTHOR(S) Thomas Jaroszewicz, Elizabeth Bleszynski, Marek Bleszynski				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Monopole Research 739 Calle Sequoia, Thousand Oaks, CA 91360			8. PERFORMING ORGANIZATION REPORT NUMBER MON- -2005-12	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research 875 North Randolph Rd., Ste 325, Room 3112 Arlington, VA 22203			10. SPONSORING / MONITORING AGENCY REPORT NUMBER AFRL-SR-AR-TR-06-0266	
11. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Air Force position, policy or decision, unless so designated by other documentation.				
12 a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited.			12 b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The objective of our effort was to develop computational methods for constructing high-frequency asymptotic solutions in scattering on perfectly conducting objects. The emphasis of the first stage of our work was to describe high frequency phenomena in terms of numerically implemented evolution of wave-fronts associated with the propagating waves. The wave-front evolution algorithm is implemented for the leading high frequency mechanisms including: free-space propagation, reflection on smooth surfaces, wave-front splitting at the shadow boundary, generation and propagation of edge diffracted wave-fronts, and surface wave propagation. The second stage of our work was directed towards developing novel fast rigorous (direct or iterative) solution methods based on construction of economical parameterization of high frequency solutions in terms of basis functions defined on large supports. The wave-front evolution technique developed in the first stage provides a numerical prescription for the selection and determination the parameters of the postulated analytical representation of such basis functions. The numerical prescription together with numerical tools constructed during this effort will constitute an important element of the planned future, high frequency solution technique employing basis functions defined on large supports.				
14. SUBJECT TERMS high frequency electromagnetic scattering, wave-front evolution, multiple scattering, creeping waves, edge diffraction			15. NUMBER OF PAGES VII+151	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OR REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION ON THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL	

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used for announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements***.

Block 1. Agency Use Only (Leave blank)

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, and volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s) project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract	PR - Project
G - Grant	TA - Task
PE - Program Element	WU - Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. (if known)

Block 11. Supplementary Notes. Enter information not included elsewhere such as; prepared in cooperation with....; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement.

Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NORFON, REL, ITAR).

DOD - See DoDD 4230.25, "Distribution Statements on Technical Documents."
DOE - See authorities.
NASA - See Handbook NHB 2200.2.
NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave Blank
DOE - Enter DOE distribution categories from the Standard Distribution for unclassified Scientific and Technical Reports
NASA - Leave Blank.
NTIS - Leave Blank.

Block 13. Abstract. Include a brief (*Maximum 200 words*) factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subject in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (NTIS *only*).

Block 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (Unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

REPORT DOCUMENTATION PAGE (SF298)
(Continuation Sheet)

MASTER COPY: PLEASE KEEP THIS "MEMORANDUM OF TRANSMITTAL" BLANK FOR REPRODUCTION PURPOSES. WHEN REPORTS ARE GENERATED UNDER THE ARO SPONSORSHIP, FORWARD A COMPLETED COPY OF THIS FORM WITH EACH REPORT SHIPMENT TO THE ARO. THIS WILL ASSURE PROPER IDENTIFICATION. NOT TO BE USED FOR INTERIM PROGRESS REPORTS; SEE PAGE 2 FOR INTERIM PROGRESS REPORT INSTRUCTIONS.

MEMORANDUM OF TRANSMITTAL

U.S. Army Research Office
ATTN: AMSRL-RO-BI (TR)
P.O. Box 12211
Research Triangle Park, NC 27709-2211

☐ Reprint (Orig + 2 copies)

☐ Technical Report (Orig + 2 copies)

☐ Manuscript (1 copy)

☐ Final Progress Report (Orig + 2 copies)

☐ Related Materials, Abstracts, Theses (1 copy)

CONTRACT/GRANT NUMBER:

REPORT TITLE:

is forwarded for your information.

SUBMITTED FOR PUBLICATION TO (applicable only if report is manuscript):

Sincerely,

Contents

1	Description of the effort	1
2	Main results	3
3	Benefits of the developed approach	6
4	Commercialization of the developed approach	7
5	Technical details: A. Scatterer surface and wavefront parameterization	10
5.1	Construction of edges	10
5.2	Construction of faces	12
5.3	Limitations of the second-order parameterization	13
5.4	Errors in second-order surface parameterization	14
5.5	Numerical examples of surface parameterization	14
6	Technical details: B. Spatial ray-tube and wavefront evolution	18
6.1	Wavefront and ray tube evolution in geometrical optics	18
6.2	Ray tube partitioning and ray interpolation	19
6.3	Ray tube partition and interpolation procedures	21
6.4	Ray tube interpolation algorithms	27
6.4.1	Construction of face boundaries and faces	28
6.4.2	Interpolation of ray directions	29
6.4.3	Interpolation of curvatures	29
6.4.4	Interpolation of fields	32
6.4.5	Computation of surface currents and the scattered field	32
6.4.6	Difficulties in ray tube interpolation procedures	32
6.5	Numerical examples of ray tube evolution	33
6.5.1	Ray tube expansion and splitting in free-space evolution	33
6.5.2	Ray tube splitting in incidence on object boundary . .	34
6.5.3	Ray tube splitting in reflections	35
6.5.4	Scattering on a system of two smooth convex objects .	36
7	Technical details: C. Evolution of creeping ray tubes (ray-strips)	44
7.1	Curved surface parameterization	44
7.2	Construction of creeping rays (geodesics)	44
7.3	Evolution of ray strips	48

7.4	Implementation of the creeping-ray construction algorithms . . .	49
7.5	Algorithms for current parameterization and computation . . .	52
7.6	Asymptotic expressions for the fields and currents due to creeping rays	52
7.7	Parameterization of the current	55
7.8	Implementation of algorithms for current computation	59
7.8.1	Computation of ray parameters	59
7.8.2	Parameterization of the current	60
7.8.3	Computation of the scattered field	60
7.8.4	Computation for multiple frequencies	60
7.9	Applications of creeping-ray algorithms	60
7.9.1	Semi-analytic construction of creeping rays for a body of revolution	64
7.9.2	Numerical ray tracing and current computation	72
8	Yale University Subcontract to Monopole Research, Topic # AF01-009	78
9	Approximate Formulae For Certain Prolate Spheroidal Wave Functions Valid for Large Values of Both Order and Band-Limit	79
9.1	Introduction	79
9.2	Mathematical Preliminaries	81
9.3	Legendre Polynomials	81
9.4	Elliptic Integrals	82
9.5	Prolate Spheroidal Wave Functions	83
9.6	Analytical Apparatus	85
9.7	Legendre Coefficients and Three-Term Recursions	85
9.8	Inverse Power Method as an Analytical Tool	87
9.9	Connections Between $\psi_m(1)$ and λ_m for Large m	88
9.10	Formulae for PSWFs and Eigenvalues χ_m, λ_m	90
9.11	Formulae for $\chi_m^c, \psi_m^c, \psi_m^c(1), \psi_m^c(0)$	91
9.12	Formulae for λ_m^c	94
9.13	Results Following From the WKB Analysis of the Equa- tion (9.2)	95
9.14	Numerical Results	97
9.15	Conclusions	99

10 A Wideband Fast Multipole Method for the Helmholtz Equation in Three Dimensions	109
10.1 Introduction	109
10.2 Mathematical Preliminaries	115
10.3 Symmetry	115
10.4 Spherical Harmonics	116
10.5 Integration of Spherical Harmonics	117
10.6 Filtering and Interpolation of Spherical Harmonics	118
10.7 Spherical Bessel and Hankel functions	120
10.8 Partial-Wave Expansions	121
10.9 Truncation Error	122
10.10 Rotation	125
10.11 Translation	126
10.12 High-Frequency Diagonal Form	127
10.13 Discretization	128
10.14 Far Field to Far Field and Local to Local Translations	129
10.15 Far Field to Local Translation	130
10.16 Exponential Expansions	133
10.17 Conversion Between Exponential and Partial-Wave Expansions	138
10.18 Algorithm	140
10.19 Numerical Results	141
10.20 Conclusions	147

List of Figures

1	A curve interpolating vertices.	11
2	Color-coded error in the piecewise-quadratic surface approximation for the ellipsoidal-type object. The boundaries of the large curved triangles are superimposed on the surface.	16
3	Color-coded error in the piecewise-quadratic surface approximation for the ogive, showing also boundaries of the curved triangles superimposed on the surface.	17
4	Errors in the quadratic surface approximation near a tip of the ogive.	17
5	A part of a wavefront divided into two ray tubes.	20
6	An example of partition of two adjacent RTs caused by RT expansion.	22
7	A schematic representation of an incident RT which undergoes partition even though all the rays intersect the scatterer.	23
8	An example of partition of two adjacent RTs due to scattering on the object boundary.	24
9	Reflection of a RT on a curved surface, increasing the RT divergence.	25
10	Partition of an incident RT, reducing divergence of the reflected RTs.	25
11	A RT incident on the scatterer surface, with interpolating rays incident at current computation points.	27
12	A procedure testing if a point $\mathbf{r}$ is contained inside the given RT.	27
13	Partitions of an expanding RT evolving in free space.	34
14	Partitions of RTs reflecting near the object boundary. The initial RTs are split into reflecting parts (a), and forward-propagating parts (b). The black curves mark the common boundary of the reflected and forward-propagating WF segments.	35
15	Partitions of RTs related to reflection from a curved surface, for the same configuration as in Fig. 14. The reflected RTs (a) result from the partition (b) of the incident RTs.	36
16	A sample RT (originating from the shown plane-wave WF) illuminating the ellipsoid S_1 , reflecting, and illuminating the ellipsoid S_2 . Shown on the ellipsoids' surfaces is the real part of the current component J_y	37

17	Single reflection contribution to the bistatic scattering cross-section on the system of two ellipsoids, for two different WF discretizations.	38
18	Single+double reflection contribution to the bistatic scattering cross-section on the system of two ellipsoids, for three different WF discretizations.	39
19	Comparison of single and single+double reflection contributions to the bistatic scattering cross-section on the system of two ellipsoids, computed for the same WF discretization. . . .	39
20	A sample RT illuminating the plate S_1 , reflecting, and illuminating the plate S_2 , for the incident WF resolution $h = 0.5$. Shown on the surfaces is the imaginary part of the current component J_z	41
21	Comparison of single and single+double reflection contributions to the bistatic scattering cross-section on the system of two flat plates, computed for three WF discretizations. . . .	42
22	Single and single+double reflection contributions, computed for the intermediate resolution ($h = 0.5$), compared with the rigorous MoM result.	42
23	Single and single+double reflection contributions (the latter computed for two resolutions, $h = 0.5$ and $h = 0.25$), compared with the rigorous MoM result, computed at twice the frequency of Figs. 21 and 22.	43
24	Behavior (in the vicinity of an ogive's tip) of a sample of creeping rays (geodesics) emerging from the other tip. . . .	47
25	An enlargement (by approximately a factor 10) of the vicinity of the ogive tip, with the set of geodesics shown in Fig. 25. .	48
26	Visualization of two nearby ray strips propagating on the surface of an ellipsoid. The heavy arrows indicate the direction of the incident plane wave.	51
27	Another view of the set of ray strips of Fig. 26, with indicated intersections of rays.	51
28	A creeping ray strip configuration used in surface current evaluation.	59
29	Currents on the spheroid (a) and cone-double-sphere (b) objects, computed with the MoM solver for the plane wave of vertical polarization incident along the negative z -axis ($\theta_i = \phi_i = 0^\circ$). The radii of the objects are approximately $R \simeq 4.54 \lambda$	61

30	The vertical- (E -plane-) polarization bistatic cross-section for the axially incident plane wave on the perfectly conducting sphere, the spheroid, and the cone-double-sphere, computed with the rigorous (MoM) solution.	62
31	The cross-section of Fig. 30 shown in a limited angular range, exhibiting different oscillatory patterns.	63
32	The vertical-polarization bistatic cross-section on the perfectly conducting sphere of $R \simeq 4.54 \lambda$, at the incidence angles $\theta_i = \phi_i = 0^\circ$: the rigorous result (Mie solution), Physical Optics (PO), and PO with creeping waves (PO+CW).	65
33	The hard current component J_t on the sphere, plotted as the function of s/λ . The absolute value of the PO current is also shown.	67
34	The vertical-polarization bistatic cross-section on the perfectly conducting prolate spheroid at 18 GHz, at the incidence angles $\theta_i = \phi_i = 0^\circ$: the rigorous result (MoM), PO, and PO with creeping waves (PO+CW).	68
35	The hard current component J_t on the spheroid, plotted as the function of s/λ , shown together with the absolute value of the PO current.	69
36	The vertical-polarization bistatic cross-section on the perfectly conducting cone-double-sphere at 18 GHz, at the incidence angles $\theta_i = \phi_i = 0^\circ$: the rigorous result (MoM), PO, and PO with creeping waves (PO+CW).	70
37	The hard current component J_t on the cone-double-sphere, plotted as the function of s/λ	72
38	The vertical-polarization bistatic cross-section on the perfectly conducting sphere at 18 GHz, at the incidence angles $\theta_i = \phi_i = 0^\circ$: the rigorous result (MoM), (PO), and numerically evaluated asymptotic solution with creeping waves (PO+CW).	73
39	The vertical-polarization bistatic cross-section on the perfectly conducting spheroid at 18 GHz, at the axial incidence: the rigorous result (MoM), PO, and numerically evaluated asymptotic solution with creeping waves (PO+CW).	74
40	The vertical-polarization bistatic cross-section on the perfectly conducting cone-double-sphere at 18 GHz, at the axial incidence: the rigorous result (MoM), PO, and the numerical computation of the asymptotic solution with creeping waves (PO+CW).	75

41	The $+z$ -list of the box b	136
42	Surface triangulation of an aircraft-shaped object 50 wavelengths in size. The size of the smallest triangle is 1.06×10^{-6} wavelengths, and the size of the largest is 2.86×10^{-1} wavelengths.	148
43	Surface triangulation of a horse 50 wavelengths in size. The size of the smallest triangle is 9.34×10^{-3} wavelengths, and the size of the largest is 3.27×10^{-1} wavelengths.	149

1 Description of the effort

The goal of the effort was the development of a computational framework for efficient application of high-frequency methods to scattering and radiation problems involving realistic geometries. A particular class of applications which motivated this work were problems involving antennas mounted on large platforms modeled by smooth curved surfaces, in the cases where the platform can be described by means of high-frequency asymptotic methods. We aimed at developing methods which would allow description of asymptotic high-frequency solutions in terms of the induced surface currents, and an *economical representation* of these currents, requiring a number of parameters growing with frequency f slower than $O(f^2)$.

The main purpose of the work carried out was to devise ray tracing and ray-tube tracing methods which would not require generating a large number of rays, comparable to the number of unknowns in the conventional Method of Moments (MoM) computation, i.e., growing proportionally to the object size in wavelength units. In order to achieve this goal, we implemented a number of algorithms which allow us to describe a wavefront as a continuous triangulated surface, consisting of **curved**, second-order triangles. During the wavefront expansion, we maintain an approximately constant resolution of the wavefront geometry by gradually increasing the number of triangles and rays. Our algorithms also adaptively adjust the number and distribution of rays in the cases when reflection of the wavefront from a curved surface causes a strong divergence of the ray tubes. Such techniques eliminate the need of generating an excessively large number of rays in the initial stage of the ray-tracing process.

Equivalently, each triangle on the wavefront surface is described as a smaller wavefront associated with a ray tube, formed by a triplet of rays. The ray-tube and wavefront evolution algorithms we implemented ensure continuity of the wavefront (or, equivalently, adjacency of the ray tubes). These algorithms are implemented for the cases of free-space wavefront propagation, reflection on smooth surfaces, and wavefront and ray tube splitting at the shadow boundary. Together with similar methods developed in Phase I for generating and evolving wavefronts due to edge diffraction, these techniques allow description of the most important high-frequency scattering mechanisms: (multiple) reflections and edge diffraction.

Further, we compute scattered fields and the resulting cross-sections by first evaluating and then integrating over the induced surface currents. In this context we developed algorithms for accurate second-order interpolation of fields within ray tubes.

The main benefit of the developed wavefront propagation method is a relatively small number of rays and ray tubes necessary to obtain an accurate description of the considered high-frequency phenomena. The required wavefront resolution (or the width of a ray tube) is not limited to a fraction of the wavelength λ , but rather (in the case of multiple reflections on smooth surfaces) is controlled by the **surface curvature radius** ρ . We found that the accuracy of 1 dB in the cross-section is, typically, achieved with the wavefront resolution h of order of $h \simeq 0.5 \rho$, practically independently of the wavelength, in the range $\rho \gtrsim 2 \lambda$.

We also applied the concepts of ray-tube evolution to modeling of propagation of creeping waves on smooth, curved, convex surfaces. In this case ray tubes become **ray strips** following the geodesics on the surface. We developed accurate numerical methods for constructing such ray strips and for computing their divergence. The latter algorithm is based on numerical solution of the geodesic deviation equation, involving the curvature of the scatterer surface.

We tested our creeping-ray algorithms and code to compute scattering cross-sections on several bodies of revolution. In this case we can compare our numerical results for ray-strip evolution on a discretized surface with the semi-analytic evaluation based on analytically known geodesics. We found a close agreement of the numerical computation and the semi-analytic results. We also found that our approach, which explicitly takes into account creeping-ray propagation in the **deep shadow region** is significantly more accurate than approximate techniques which only consider the shadow transition region and extrapolate the creeping-ray contributions based on the local curvature information at the shadow boundary.

Our main results are briefly summarized in Section 2, and the technical details are described in Sections 5 – 7.

The activities of the Yale University group are summarized in Section 8, followed by two additional Sections.

1. The first part of the report (Section 9) is concerned with mathematical tools useful in evaluating prolate spheroidal wave functions. Such band-limited functions may arise in constructing large-support basis functions (or directional basis functions) as an alternative parameterization of high-frequency solutions in scattering problems.
2. The second part of the report (Section 10) concentrates on development of a unified Fast Multipole Method applicable to physical systems containing both electrically large portions, described by typical high-frequency discretizations, and parts characterized by highly

sub-wavelength details. Such situations occur in problems involving antennas on large platforms.

2 Main results

A. The initial emphasis of our work under this contract was to develop computational methods for constructing high-frequency asymptotic solutions in scattering on perfectly conducting objects. Our general approach was to describe high-frequency phenomena in terms of numerically implemented evolution of the wavefronts and ray tubes associated with the propagating waves.

We put much effort in developing methods applicable to scattering on curved surfaces. Typical ray-tracing techniques are based on independent evolution of rays and do not involve the concept of a wavefront as a numerically modeled surface. As a result they encounter difficulties in describing wave reflection from curved surfaces, due to the divergence of rays and the resulting lack of information on the wavefront geometry.

We avoid these difficulties by relying on the wavefront concept, and implementing it numerically as a well-defined geometrical surface.

In the area of modeling wavefront and ray-tube evolution we developed algorithms and computational tools for describing:

1. Free-space evolution of wavefronts in Geometric Optics (GO):
 - a. Modeling of wavefronts as continuous discretized (triangulated) surfaces, with triangles associated with ray tubes.
 - b. Modeling of wavefront evolution in terms of evolution of a contiguous set of ray tubes. Divergent ray tubes are automatically split into narrower tubes to ensure an approximately constant density of rays, and a sufficient resolution in the wavefront definition.
 - c. Piecewise quadratic parameterization and interpolation of the wavefront and their segment, fully taking into account the wavefront curvature.
2. Multiple reflection of wavefronts on curved surfaces, described in Geometrical Optics:
 - a. Fast and accurate evaluation of intersections of rays with a curved surface.

- b. Computation of the curvatures of reflected ray tubes and the wavefront based on the curvature of the incident wavefront and the curvature of the scatterer surface.
 - c. Adaptive ray tube splitting reducing divergence angles of the reflected ray tubes, and ensuring sufficient resolution of the wavefront. Ray-tube splitting is controlled by the curvature of the scatterer surface and the incidence angle of the ray tube.
 - d. Quadratic interpolation techniques in reflected ray tube construction and splitting, ensuring continuity of the wave front, i.e., matching of boundaries of adjacent ray tubes.
- 3. Evaluation of currents induced by fields associated with wavefronts and ray tubes:
 - a. Fast computation of intersections of ray tubes with the scatterer surface, taking into account scatterer surface and wavefront curvatures.
 - b. Accurate second-order interpolation of fields within ray tubes.
- 4. Numerical modeling of creeping waves propagating on convex surfaces (smooth or with sharp edges) as “ray strips” – surface analogues of ray tubes:
 - a. Techniques for describing evolution of creeping rays as geodesics on general convex piecewise smooth surfaces.
 - b. A novel procedure efficient computing of the change in the ray strip width, based on a numerical solution of the geodesic deviation equation. This method takes into account effects of ray strip traversing edges on the surface (formally, lines of infinite curvature).
 - c. Algorithms for evaluating currents associated with creeping rays, taking into account their polarization- and curvature-dependent attenuation.

We demonstrated the developed methods in computation of currents and scattering cross-sections for a number of high-frequency processes involving convex smooth curved-surface objects and sets of such objects. Particular cases included multiple reflections on a system of two smooth convex objects,

and propagation of creeping waves in the shadow region on a single convex object.

As technical tools, we developed codes and code modules implementing algorithms for:

- Piecewise second-order parameterization of curved scatterer surfaces, with computation of all standard surface characteristics, such as curvatures, curvature tensor, and Christoffel symbols.
- A similar piecewise second-order parameterization of the wavefront surface.
- Fast location of intersections of rays and ray tubes with the scatterer surface.
- Second-order interpolation of ray-tube fields, as required in ray-tube splitting and induced current computation.
- Parameterization of creeping rays and their divergence in terms of geodesic deviation equation.

B. In the later stage of the work we directed part of the effort towards methods applicable in developing novel fast rigorous (direct or iterative) solution methods. We were considering solution methods based on construction of economical parameterization of the high-frequency solutions (currents) in terms of basis functions (a) defined on large supports (of at least several wavelengths size), (b) characterized by collimated radiation patterns, and (c) constructed numerically. We refer to these functions as “numerical” directional basis functions (DBFs). Their particularly important property is the angular collimation of the radiated fields, which leads to sparsity of the impedance matrix in the DBF representation.

In the “numerical” DBF approach the basis functions are not constructed as analytic representations of predicted high-frequency behavior of the solutions, but rather as eigenfunctions of a certain integral operator related to the power radiated by all possible source (current) distributions on the scatterer surface.

The two main advantages of the numerically constructed DBFs are (1) their versatility and applicability to complex geometries and (2) a better collimation of the radiation patterns, and hence higher degree of resulting impedance matrix sparsity.

However, at present, the drawback of the approach is a high computational cost of DBF construction. The DBF construction algorithms are a subject of active research, which may result in total cost scaling, up to logarithms, as the number of MoM unknowns.

The large problem size in the *direct-solution* approach utilizing DBFs is related to the fact that full set of DBFs has to be constructed such as to parameterize all possible solutions (current distributions) for all possible incident waves. An alternative approach would be an *iterative solution*, applicable to a single incident wave or a limited set of incident waves. In this case one would construct only a subset of the DBFs, relevant to the considered excitation(s), reducing in this way the problem complexity. Our wavefront and ray-tube evolution techniques provide, in this context, an efficient way of selecting those DBFs which strongly couple to the expected solution. In order to generate a sufficiently large set of DBFs, it is important to keep track of all relevant scattering processes, and hence to generate complete sets of ray tubes and wavefronts. On the other hand, the accuracy of the asymptotic theories (GDT, UTD) is of secondary importance, since they are used only to estimate the significance of the individual processes, and not for detailed predictions of the fields and currents.

3 Benefits of the developed approach

The main advantages of the approach we developed are as follows:

1. In the area of methods for spatial ray-tube and wavefront evolution, involving free-space wavefront propagation and reflection on flat-faceted and curved surfaces, we developed techniques allowing using a number of rays N_r growing with the frequency much more slowly than in conventional ray tracing methods requiring spacing between rays scaling proportionally to the wavelength. By using ray-tube partition techniques, together with second-order methods for scatterer surface and wavefront parameterization and interpolation, we were able to reduce the scaling in the number of rays with frequency f to $N_r \sim f^{2/3}$ compared to $N_r \sim f^2$ in the conventional approach.
2. In the area of problems involving creeping waves propagating on smooth convex surfaces, we developed highly accurate techniques for computing evolution of ray strips (the analogue of ray tubes). These techniques allow an accurate computation of geodesics, evaluation of ray-strip spreading, and, eventually, accurate numerical evaluation of

the asymptotic theory predictions for the currents in the deep-shadow region. We have shown that computations taking into account the actual asymptotic behavior of currents in the shadow region reproduce effects that are missing in approximate methods based on extrapolation of the solutions from the shadow boundary.

3. We developed an extensive set of software modules implementing the techniques mentioned in the points 1 and 2 above. These modules are available in the form of a library of modules, and two main programs, for modeling of scattering processes involving spatial ray-tube evolution and evolution of creeping rays.
4. The methods we developed can be applied both to *direct computation* of asymptotic high-frequency solutions, and as a *auxiliary tool* in numerical construction of systems of “*directional basis functions*” (*DBFs*), characterized by large supports and narrow radiation patterns. In the latter case wavefront and ray-tube evolution methods allow pre-selection of the relevant DBFs, corresponding to dominant scattering mechanisms in the given scattering problem.

4 Commercialization of the developed approach

The software tools developed during Phase I and II offer the following features:

- A high-frequency solver based on wavefront evolution methods. It uses as input the scatterer geometry and a definition of the initial wavefront (radiation source), and generates as output evolved wavefronts, induced surface currents, and radiated fields.
- The models, in a consistent way, multiple reflection processes and the most important diffraction phenomena (edge and smooth-surface diffraction). The structure of the code allows later inclusion of other diffraction mechanisms, such as crack and gap diffraction, as their UTD descriptions become available.
- The solver is applicable to perfectly conducting scatterers described either by facetized representations with flat facets, or by parametric representations.
- The solver offers trade-offs between the accuracy and the computational cost by allowing the user to include or exclude specific scattering mechanisms.

- The software package includes basic tools for geometry manipulation and preprocessing, as well as essential graphical tools for visualization of the geometry, wavefronts, and induced currents.
- The solver and auxiliary programs are provided in the form of modular function libraries, to facilitate their interfacing with other software. In particular, they will be amenable to use in the context of hybrid methods, and methods based on asymptotic integral-equation formulations.
- The solver is portable to Unix, Linux, and PC platforms.

The algorithms developed under this contract will form a backbone for the commercial version of the high frequency electromagnetic simulation software to be subsequently developed by Monopole Research. The software will be packaged for

- (i) defense applications: performance analysis of radio frequency (RF) antennas installed and operating in the presence of realistic platforms, far field and near field radar signature predictions, RF imaging), and
- (ii) commercial applications: urban propagation, remote sensing of vehicles.

There are over 500 organizations across the country using EM simulation codes to produce and analyze scattering data for realistic aircraft, missiles, ships, spacecraft, and ground vehicles. They include numerous government organizations and major aerospace firms such as Lockheed Martin, Northrop Grumman, Boeing, Sikorsky Aircraft, Raytheon, and TRW.

There are also several Government organizations sponsoring developments of novel solvers for realistic EM simulations; they include: the Defense Advanced Research Projects Agency (DARPA), the Air Force Research Lab (AFRL), the Air Force Office of Scientific Research (AFOSR), the Army Research Lab (ARL), the Naval Research Lab (NRL), the Office of Naval Research (ONR), Sensors, Space Vehicles or Human Effectiveness Directorates, the Office of the Secretary of Defense/High Performance Computing (OSD/HPC), the National Air Intelligence Center (NAIC), the Missile and Space Intelligence Center (MSIC), the Naval Air Warfare Center (NAWC).

We intend to continue contacts with potential Government and commercial customers, to identify their specific needs and to include in the code development the capabilities of interest to prospective customers. Recently the following companies have contacted us with respect to the future evaluation/purchasing of the high frequency simulation software: TRW (software for antenna applications involving wide-band pulses), Titan Corporation

(modeling of antennas with dispersive surface parameters), Rockwell Scientific (large scale computations of phase-array), Boeing Autometric (simulation of urban propagation problems).

5 Technical details: A. Scatterer surface and wavefront parameterization

We briefly summarize here the algorithms we developed and implemented in order to describe curved surfaces as collections of second-order parametric curved triangles. They were originally applied to wavefronts composed out of wavefront segments associated with the individual ray tubes (in which case only the second-order parameterization is available), but we found them useful also in parameterizing smooth convex scatterer surfaces.

Our second-order parameterization uses as input vertices on the surface (i.e., vertices of the curved triangles), and normals to the surface at the vertex locations. This information allows us to determine uniquely the curved (second-order) edges of the triangles, and the corresponding second-order parameterization of the triangles themselves.

The details of the construction are described, for the general case, in Appendix A of Report 1, and in Appendices B and C of that report for the case of small-curvature triangles (or for ray tubes of small angular widths). We summarize here the main points.

5.1 Construction of edges

We first consider two vertices, $\mathbf{r}_1$ and $\mathbf{r}_2$, and the corresponding normal directions $\mathbf{n}_1$ and $\mathbf{n}_2$ (Fig. 1). On this basis we construct the edge connecting the vertices as a parametric curve $\mathbf{r}(t)$, $t \in [0, 1]$, and a normal field $\mathbf{n}(t)$ such that

$$\mathbf{r}(0) = \mathbf{r}_1 \ , \quad \mathbf{r}(1) = \mathbf{r}_2 \ , \quad (5.1)$$

$$\mathbf{n}(0) = \mathbf{n}_1 \ , \quad \mathbf{n}(1) = \mathbf{n}_2 \ . \quad (5.2)$$

Obviously, we also have to require that the normal field is orthogonal to the tangent to the curve, i.e.,

$$\dot{\mathbf{r}}(t) \cdot \mathbf{n}(t) = 0 \ , \quad (5.3)$$

where $\dot{\mathbf{r}}$ is the derivative of $\mathbf{r}$.

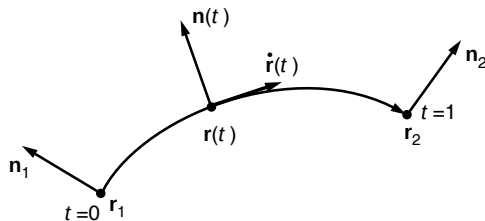


Figure 1: A curve interpolating vertices.

We construct the edge using the lowest order, quadratic parameterization of the curve, and a close-to-linear parameterization of the normal field (the latter is relevant only in applications to wavefronts).

Since a quadratic curve is planar, we have to define its plane, say P_{12} . If the normal $\mathbf{n}_1$ and $\mathbf{n}_2$ are coplanar they uniquely specify P_{12} . Otherwise, we define P_{12} , by means of a least-squares procedure, to lie as closely as possible to the normal $\mathbf{n}_1$ and $\mathbf{n}_2$.

More specifically, we define the curve as

$$\mathbf{r}(t) = \mathbf{r}_1 (1 - t) + \mathbf{r}_2 t + \mathbf{B} t (1 - t) , \quad (5.4)$$

and parameterize $\mathbf{B}$ as

$$\mathbf{B}(\gamma) = \beta_1 \mathbf{n}_1 + \beta_2 \mathbf{n}_2 + \gamma \mathbf{m} , \quad (5.5)$$

with

$$\mathbf{m} = \mathbf{D}_1 + \mathbf{D}_2 , \quad (5.6)$$

where

$$\mathbf{D}_1 = \mathbf{R} \times \mathbf{n}_1 , \quad \mathbf{D}_2 = \mathbf{R} \times \mathbf{n}_2 , \quad (5.7)$$

are normal to the planes specified by the ray directions $\mathbf{n}_i$ and the edge $\mathbf{R}$

$$\mathbf{R} = \mathbf{r}_1 - \mathbf{r}_2 . \quad (5.8)$$

To determine the parameters β_i and γ we now require, according to Eq. (5.3), that the tangent to the curve,

$$\dot{\mathbf{r}}(t) = -\mathbf{R} + \mathbf{B} (1 - 2t) \quad (5.9)$$

is orthogonal (at $t = 0$ and $t = 1$) to the normal vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ (Eq. (5.3)), i.e.,

$$\mathbf{n}_1 \cdot \mathbf{B} = \mathbf{n}_1 \cdot \mathbf{R} , \quad \mathbf{n}_2 \cdot \mathbf{B} = -\mathbf{n}_2 \cdot \mathbf{R} . \quad (5.10)$$

These conditions specify a one-parameter set of vector $\mathbf{B}$, determined by the intersection of the planes defined by Eq.(5.10). The remaining parameter is found by minimizing the least-squares error

$$\Phi = (\mathbf{D}_1 \cdot \mathbf{B})^2 + (\mathbf{D}_2 \cdot \mathbf{B})^2 . \quad (5.11)$$

As a result, we obtain a set of linear equations for the coefficients β_1 , β_2 , and γ .

In the case of $(\mathbf{n}_1 \cdot \mathbf{n}_2)^2 \neq 1$ the coefficients β_i are given by

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \frac{1}{1 - (\mathbf{n}_1 \cdot \mathbf{n}_2)^2} \begin{bmatrix} 1 & -\mathbf{n}_1 \cdot \mathbf{n}_2 \\ -\mathbf{n}_1 \cdot \mathbf{n}_2 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{n}_1 \cdot \mathbf{R} \\ -\mathbf{n}_2 \cdot \mathbf{R} \end{bmatrix} , \quad (5.12)$$

and γ is found from the equation

$$\beta_2 \mathbf{D}_1 \cdot \mathbf{n}_2 \mathbf{D}_1 \cdot \mathbf{m} + \beta_1 \mathbf{D}_2 \cdot \mathbf{n}_1 \mathbf{D}_2 \cdot \mathbf{m} + [(\mathbf{D}_1 \cdot \mathbf{m})^2 + (\mathbf{D}_2 \cdot \mathbf{m})^2] \gamma = 0 . \quad (5.13)$$

If the vectors $\mathbf{R}$, $\mathbf{n}_1$, and $\mathbf{n}_2$ are coplanar, $\mathbf{D}_1 \cdot \mathbf{n}_2 = \mathbf{D}_2 \cdot \mathbf{n}_1 = 0$, and $\gamma = 0$.

In the case of $(\mathbf{n}_1 \cdot \mathbf{n}_2)^2 = 1$ the only possible configuration for a convex surface is $\mathbf{n}_1 = \mathbf{n}_2$ and $\mathbf{n}_1 \cdot \mathbf{R} = \mathbf{n}_2 \cdot \mathbf{R} = 0$, representing a flat surface; hence a geometrically justified solution is $\mathbf{B} = 0$.

Given the boundary curve $\mathbf{r}(t)$, we construct now the normal field $\mathbf{n}(t)$ as a linear interpolation modified by a small correction term,

$$\mathbf{n}(t) = \mathbf{n}_1 (1 - t) + \mathbf{n}_2 t - f(t) [\mathbf{R} - \mathbf{B} (1 - 2t)] , \quad (5.14)$$

where, as follows from Eqs. (5.3) and (5.9),

$$f(t) = 2 \frac{\mathbf{R} \cdot (\mathbf{n}_1 + \mathbf{n}_2) t (1 - t)}{[\mathbf{R} - \mathbf{B} (1 - 2t)]^2} . \quad (5.15)$$

5.2 Construction of faces

Having determined curved edges, we construct faces by means of second-order surface interpolation, i.e., we define curved second-order faces having as boundaries the quadratic curves parameterized as in Eq. (5.4). In terms

of the barycentric coordinates $u^1 \geq 0$, $u^2 \geq 0$, and $u^3 \geq 0$, satisfying $u^1 + u^2 + u^3 = 1$, the surface can be represented as

$$\begin{aligned} \mathbf{r}(u^1, u^2, u^3) = & \mathbf{r}_1 u^1 + \mathbf{r}_2 u^2 + \mathbf{r}_3 u^3 \\ & + \mathbf{b}_1 u^1 (1 - u^1) + \mathbf{b}_2 u^2 (1 - u^2) + \mathbf{b}_3 u^3 (1 - u^3) , \end{aligned} \quad (5.16)$$

where

$$\mathbf{b}_1 = \frac{1}{2} (-\mathbf{B}_1 + \mathbf{B}_2 + \mathbf{B}_3) , \quad (5.17)$$

$$\mathbf{b}_2 = \frac{1}{2} (-\mathbf{B}_2 + \mathbf{B}_3 + \mathbf{B}_1) , \quad (5.18)$$

$$\mathbf{b}_3 = \frac{1}{2} (-\mathbf{B}_3 + \mathbf{B}_1 + \mathbf{B}_2) , \quad (5.19)$$

and the vectors $\mathbf{B}_i$ are determined as described in the previous Section. Here $\mathbf{B}_2$ is associated with the boundary curve $(\mathbf{r}_1, \mathbf{r}_2)$, and similarly for other vectors, by cyclic permutations. Relations converse to Eqs. (5.17), (5.18), and (5.19) are

$$\mathbf{B}_1 = \mathbf{b}_2 + \mathbf{b}_3 , \quad (5.20)$$

$$\mathbf{B}_2 = \mathbf{b}_3 + \mathbf{b}_1 , \quad (5.21)$$

$$\mathbf{B}_3 = \mathbf{b}_1 + \mathbf{b}_2 . \quad (5.22)$$

5.3 Limitations of the second-order parameterization

As follows from the nature of the second-order surface parameterization, it is directly applicable to convex surfaces only: in order to describe a non-convex surface near its inflection points, at least a cubic parameterization would be necessary.

The second-order parameterization described above may also be, for some purposes, not fully satisfactory, since the obtained interpolation is only C^0 continuous along the boundaries of faces. The normals and curvatures are not continuous, i.e., the normals computed in two-dimensional interpolation on the face may not precisely converge to normals computed on an adjacent face, or determined from one-dimensional interpolation along the edge.

G^1 or G^2 continuity (i.e., continuity of the normals or curvature) can be achieved by utilizing higher-order (typically, rational) surface interpolations. In particular, a G^1 -continuous surface can be constructed by using the algorithm introduced by Nielsen [3]. The method uses as input three curved edges together with the specified normal fields, and constructs a curved face spanned by the edges and having the normal approaching the prescribed

normal along the edges. This procedure ensures thus that the normals of two faces joined along the edge are continuous. Similarly, G^2 -continuous faces can be generated using the method of Ref. [4].

In practice, however, the discontinuities are small, especially since our construction criteria ensure that curvature radii ρ are large compared to the transverse RT sizes h (typically, $h < .2\rho$), and discontinuities are of second order in h/ρ .

5.4 Errors in second-order surface parameterization

In applications to scatterer surface parameterization, we partition the surface into relatively large curved triangles, and use piecewise second-order parameterizations of the individual triangles. The sizes of the curved triangles are dictated by the surface curvature, and by the wavelength in the considered scattering problem.

The errors in the surface definition due to the quadratic approximation are of the order $e \sim h^3/\rho^2$, where h is the triangle size and ρ the surface curvature. In order for this approximation to be adequate, we have to require that the error in the phase

$$\delta = \frac{e}{\lambda} \sim \frac{h^3}{\rho^2 \lambda} , \quad (5.23)$$

is sufficiently small ($\delta \ll 1$). This condition imposes a limitation on the triangle size,

$$h \lesssim (\delta \lambda \rho^2)^{1/3} , \quad (5.24)$$

i.e., the triangle size has to decrease (although rather slowly) with the increasing frequency. For a more general surface parameterization of the n -th order ($n = 2$ for the quadratic parameterization), the restriction on the triangle size is

$$h \lesssim (\delta \lambda \rho^n)^{1/(n+1)} . \quad (5.25)$$

5.5 Numerical examples of surface parameterization

We have tested the interpolation procedure on a number of examples of smooth convex objects, including ellipsoids and ellipsoid-like surfaces, as well as ogives. We found that the accuracy of the interpolation strongly depends on the accuracy with which the normals are determined. That sensitivity is particularly high in regions in which the surface curvature

varies rapidly (although the curvature itself may be relatively small). The problem stems from the fact that, as follows from Eq. (5.12) the components of the curvature parameter $\mathbf{B}$ along the edge direction may be relatively large, even if the surface curvature is small:

$$\hat{\mathbf{R}} \cdot \mathbf{B} \simeq R \frac{(\mathbf{n}_1 + \mathbf{n}_2) \cdot \hat{\mathbf{R}}}{(\mathbf{n}_1 - \mathbf{n}_2) \cdot \hat{\mathbf{R}}}, \quad (5.26)$$

where we assumed coplanar normals. Hence, as long as $\mathbf{n}_1 \cdot \mathbf{R} > 0$ and $\mathbf{n}_2 \cdot \mathbf{R} < 0$ (which is always the case for a convex surface), we have

$$|\hat{\mathbf{R}} \cdot \mathbf{B}| \leq R, \quad (5.27)$$

but the component of $\mathbf{B}$ along $\mathbf{R}$ may remain finite even in the limit $\mathbf{n}_i \cdot \mathbf{R} \rightarrow 0$, when the normal components of $\mathbf{B}$ and the edge curvature $\kappa \simeq (\mathbf{n}_1 - \mathbf{n}_2) \cdot \hat{\mathbf{R}}/R$ become small.

In view of the strong dependence of the surface parameterization on the normals, the normal vectors should be evaluated, whenever possible, by using the exact surface parameterization. If such a parameterization is not available, we use the finest available triangulation of the surface. Only as the last resort would we use the coarse triangulation on which the curved triangle parameterizations are based.

If the normals are to be determined from a surface discretization with flat triangles, we evaluate them, for each vertex, as weighted averages of normals of triangles sharing the considered vertex (with the weights proportional to the angles spanned by the edges of the triangles). We note, however, that any procedure of computing normals from a surface triangulation may be unreliable: in particular, a triangulation of smooth convex surface may not even be convex.

We illustrate now our surface interpolation procedure with some examples.

As the first case, we consider a slightly irregular ellipsoidal-type object, shown in Fig. 2, of size $12 \times 8 \times 3.2 \text{ m}^3$. The original surface representation was obtained from a CAD program in the form of an IGES file. As the basis of the interpolation procedure we have assumed a fine-resolution discretization of that surface, in the form of a triangulation with 135,736 flat facets. This discretization was used to determine the normals to the surface in our interpolation procedure, in which we partitioned the object surface into 376 rather large triangles with the average edge length about 7.5 cm.

Fig. 2 shows the surface with color-coded errors in the surface interpolation, defined as the deviation between the original (finely discretized) and

the interpolated surface. It is seen that the largest deviations seem to occur in the areas where the surface curvature changes most rapidly, consistently with the fact that the quadratic parameterization is in these cases most sensitive to the definition of normals. The maximum deviation is about 1.7 cm and the average deviation only 2.6 mm.

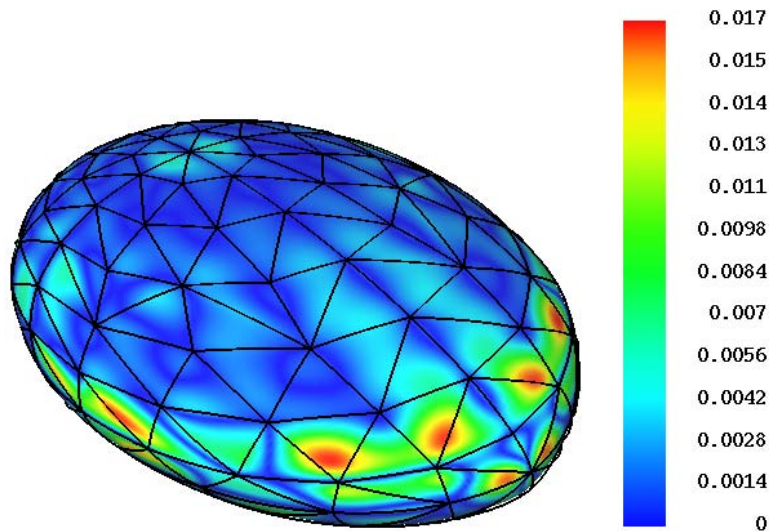


Figure 2: Color-coded error in the piecewise-quadratic surface approximation for the ellipsoidal-type object. The boundaries of the large curved triangles are superimposed on the surface.

As the second example, we show in Fig. 3 interpolation errors for an ogive of length 10 m and diameter 2 m, with the surface interpolated using 800 curved triangles of average edge length 38 cm. In this case we computed normals using the exact surface parameterization.

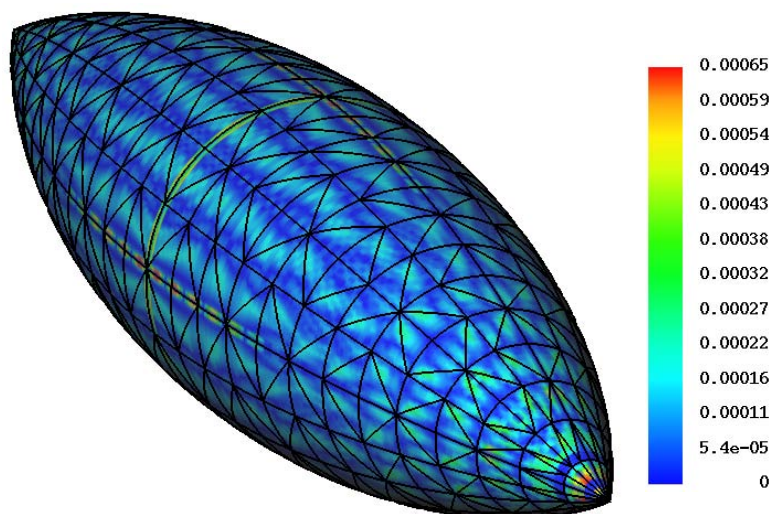


Figure 3: Color-coded error in the piecewise-quadratic surface approximation for the ogive, showing also boundaries of the curved triangles superimposed on the surface.

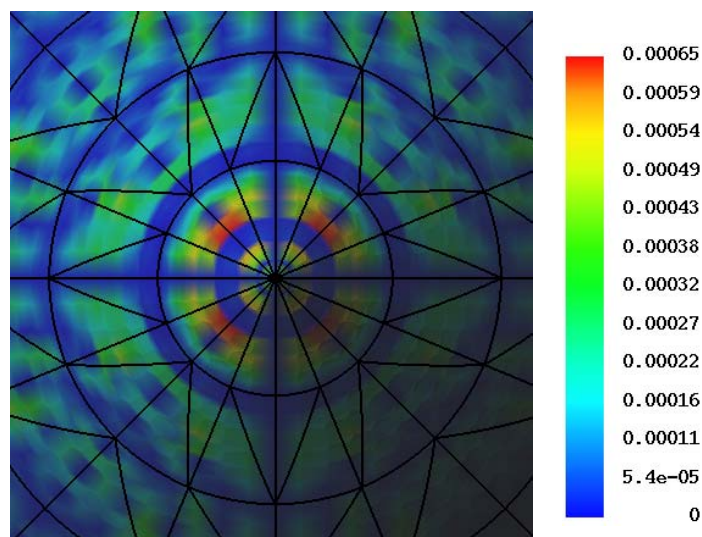


Figure 4: Errors in the quadratic surface approximation near a tip of the ogive.

Deviations between the original and approximated surface are most visible near the ogive’s tips, and along the junctions between the geometry segments (as the “exact” geometry was constructed by joining reflections, in the three symmetry planes, of the 1/8-th of the object). The latter deviations are evidently an artifact of the slight inaccuracies of the CAD program we used. The interpolation errors near the tip (Fig. 4) are more significant. However, the maximum error is only 0.65 mm, and the average error is 0.065 mm.

6 Technical details: B. Spatial ray-tube and wavefront evolution

Our general approach is to compute evolution of a wavefront incident on a scatterer by using geometrical optics (GO), and Uniform Geometrical Theory of Diffraction (UTD), to evaluate and parameterize in an efficient way currents induced on the scatterer’s surface, and to use the currents to compute the scattered fields.

Our work concentrated on scattering on smooth objects in the scope of Geometrical Optics, and in the context of creeping waves (as described in Sec. 7). Since our algorithms use sparsely distributed rays, they necessitate development of accurate algorithms for **ray interpolation**. Numerical implementation of these algorithms was a major part of code development.

6.1 Wavefront and ray tube evolution in geometrical optics

We define a wavefront (WF) as a surface of constant phase (evolution parameter s) of the field, orthogonal to a set of rays. For a given value of the evolution parameter s (defined as the optical patch length) we implement the WF numerically as a discretized triangulated surface defined by a set of points $\mathbf{r}_i$, and a set of unit ray directions $\mathbf{n}_i$.

We consider a ray, for a given evolution parameter s , as specified by the position (*ray vertex*) $\mathbf{r}$, the direction $\mathbf{n}$, two principal curvatures κ_1 and κ_2 (we assume $|\kappa_1| \geq |\kappa_2|$), and the principal curvature direction $\mathbf{a} \equiv \mathbf{a}_1$ associated with the first curvature (the other principal curvature direction is then given by $\mathbf{a}_2 = \mathbf{n} \times \mathbf{a}_1$).

The geometric optics (GO) evolution of the WF is determined by the evolution of the rays (ray tracing). Our general strategy was to implement WF evolution in terms of evolution of *ray tubes* (RTs), defined as triplets of rays whose positions form a triangle of the WF surface (Fig. 5). We

formulated the evolution algorithm in such a way that the individual RTs can be evolved ***independently***. There are several advantages of this approach compared to the algorithm based on the simultaneous evolution of all rays constituting the WF:

1. Evolution of separate RTs allows us to relax the restrictions on the length of the evolution steps. A RT may be evolved in steps whose length is controlled only by
 - (a) the transverse spreading of the RT, which requires RT subdivision, and
 - (b) scattering (reflection or diffraction) events, controlled by the intersections of the RT with the scatterer geometry.
2. Separate evolution of RTs leads to a significant simplification of the WF evolution algorithm.
3. Independent RT evolution leads to an algorithm that allows straightforward parallelization.

We stress that the independent evolution of RTs does not in any way compromise the generality and accuracy of the approach based on the evolution of a WF as a whole. The only feature of the original WF evolution scheme which cannot be implemented by means of independent RT evolution is a *reduction* of the number of rays by merging several RTs into a single RT in the case when transverse sizes of the RT decrease. In practice, however, this does not seem to be a serious limitation, since the RTs typically expand rather than contract. In any case, the lack of a mechanism for reducing the number of RTs does not affect the accuracy of the algorithm, and can only lead to a minor increase in the computational cost.

6.2 Ray tube partitioning and ray interpolation

RT evolution in GO involves free-space propagation and reflection from the scatterer. In order to ensure an adequate accuracy of ray tracing, we impose two conditions on the RTs:

- (A) A RT should not be too divergent, i.e., the maximum angle between the ray should not exceed some predefined values α_0 . This condition is necessary to ensure sufficient accuracy of WF interpolation.

- (B) A RT should not be too wide, i.e, the maximum transverse distance between the rays should not exceed some specified value h_0 . The purpose of this condition is to achieve sufficient accuracy in resolution of scatterer geometry details.

These conditions imply the necessity of **partitioning** RTs in the course of their evolution. Other processes that lead to RT partition and to creation of new RTs are reflection events (which we discuss later in this report) and diffraction.

In these RTs subdivision processes it is necessary to create new rays that will form (together with some original rays) new RTs. To these new rays have to assign their attributes: position, propagation direction, principal curvatures and curvature directions, and field values. This requires **interpolation** of the ray properties within the RT.

Another situation when ray interpolation is required is computation of currents induced on the scatterer surface. In this case we create new rays intersecting the surface of the scatterer at prescribed points.

As we describe below, we have been able to implement these interpolation algorithms in such a way that, although they require information associated with **one RT only**, they respect the **consistency of evolution** relating adjacent RTs. We discuss the features of this implementation in the following Sections.

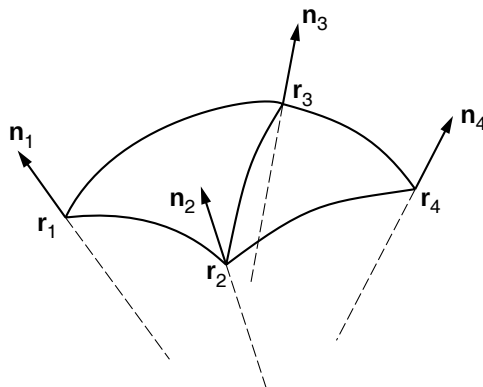


Figure 5: A part of a wavefront divided into two ray tubes.

6.3 Ray tube partition and interpolation procedures

Our algorithm implementation handles GO scattering on smooth objects. In practice, the restriction on the geometry and the RT features is that the transverse RT size (denote h) must be sufficiently small compared to the curvature radii ρ_S of the scatterer surface. We find on the basis of numerical experiments that a good accuracy can be achieved for $h \lesssim \rho_S/5$.

The implemented elements of the algorithm are as follows:

- (i) Partitioning of an expanding RT, in order to maintain its bounded transverse size.
- (ii) Computation of RT intersection with the scattering object. We distinguish here three cases, in which one, two, or three rays intersect the scatterer surface. In the latter case the entire RT undergoes reflection (although, as we describe later, it may be subdivided). In the first two cases, the RT is partitioned into a reflected part and a “forward propagating” part.
- (iii) The above-mentioned partitioning of the RT partly intersecting the object.
- (iv) Partitioning of RTs *prior* to reflection on a curved surface, in order to limit the angular divergence of the reflected RT.
- (v) Computation of surface currents using fields associated with rays intersecting the surface.
- (vi) “One-dimensional” interpolation of rays within a RT along the RT edges, required in the partitioning procedures (i), (iii), and (iv).
- (vii) “Two-dimensional” interpolation of rays within the transverse section of a RT, required in computation of currents.

We now briefly describe the first five algorithm elements. We will discuss interpolation in the next Section.

(i) To handle RT transverse expansion we consider its face (formed by the points on the three rays), and simply divide into two parts that side of the face which exceeds the prescribed transverse size h_0 (Fig. 6). We then create (by interpolation) a new ray emerging from the partition point $\mathbf{p}$, and partition the original face into two. This process is illustrated in Fig. 6 for two adjacent RTs (original faces T_1 and T_2). As seen in Fig. 6, this algorithm can be applied to each of the RT ***independently***, provided it results

in the **same** new interpolated ray starting from the partition point $\mathbf{p}$. Our presently implemented interpolation algorithm ensures the same positions and directions of the rays generated within each of the two RTs; the curvatures and fields assigned to the two rays are, however, only approximately the same. We discuss this problem in more detail when we describe the interpolation procedure in the next Section.

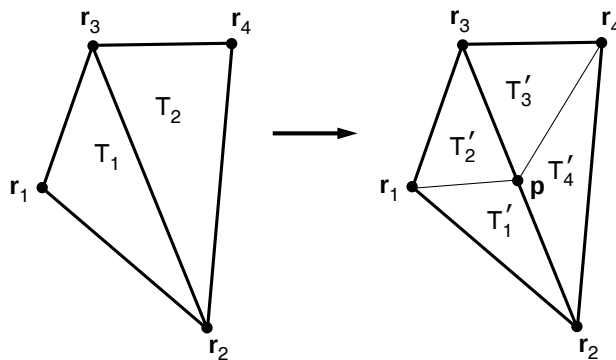


Figure 6: An example of partition of two adjacent RTs caused by RT expansion.

(ii) Intersections of the RT with the scattering object are classified on the basis of the intersections of its constituent rays; whenever an intersection is found, we store information on the intersection event: the intersection point, surface normal at this point, and surface curvature.

If all or none of the rays intersect the object, then the RT is reflected as a whole or is not reflected at all. In the remaining cases it is split into parts which undergo reflection, and which propagate without interaction. We describe the partition procedure in the next point. Here we only comment on the errors associated with the partition: Since the RT is divided by a straight line, we do not take into account the curvature of the scatterer boundary. If the transverse size of the RT is of order h_0 , and the curvature radius of the scatterer boundary is R , the *relative* error in the TR partitioning is of order h_0/R .

We also mention that implementation of the intersection algorithm requires some care in correctly identifying intersections of the RT rays with various parts of the scatterer. In particular, we have to distinguish situations when the rays intersect a contiguous segment of the scatterer surface, and situations where, for instance, all the rays intersect the surface, but their

footprint is not a connected area. An example of the latter configuration is shown schematically in Fig. 7. Rays 1 and 2 shown there span a contiguous footprint on the scatterer S_1 surface, while ray 3 intersects a separate scatterer part S_2 . Thus, the RT should be partitioned by interpolating a new ray between the rays 2 and 3; then, one part of the RT should be reflected from S_1 and the other part from S_2 .

At the moment we have no fully satisfactory algorithm able to distinguish such and similar cases, and, in fact, construction of such an algorithm is possible only within the geometry resolution limitations imposed by the finite transverse size of the RTs.

In the present implementation we partially solved the problem by identifying both “entering” and “exiting” ray intersections (Fig. 7), and imposing conditions on the overlap of intervals between these pairs of intersections. We also distinguish various intersection cases by analyzing intersection points and angles between the incident rays and the surface normals, in order to determine if the RT could have intersected a contiguous surface segment.

In general, identification of intersection cases will have to involve a precise formulation of the conditions the geometry should satisfy in order to avoid ambiguities, and will require devising methods for automatic verification of these conditions.

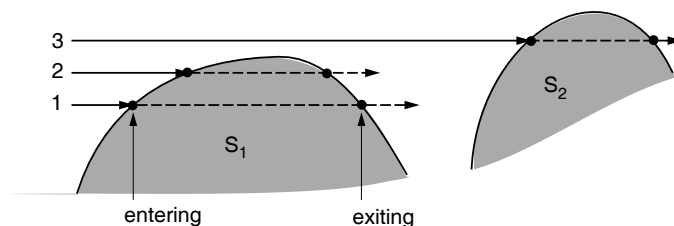


Figure 7: A schematic representation of an incident RT which undergoes partition even though all the rays intersect the scatterer.

(iii) A typical situation occurring for RTs incident near the boundary of the scatterer is as illustrated in Fig. 8. We show there two adjacent RTs such that for each of them one ray (ray of position $\mathbf{r}_2$) intersects the object, and the other two do not. Our algorithm, when applied to the RT T_1 , finds two interpolating rays passing arbitrarily closely to the object boundary. These new rays start at the points $\mathbf{p}_1$ and $\mathbf{p}_2$ obtained by one-dimensional interpolation on the edges $(\mathbf{r}_1, \mathbf{r}_2)$ and $(\mathbf{r}_3, \mathbf{r}_2)$. The original and the new rays

are then used to create a new triangulation (T'_1, T'_2, T'_3) of the subdivided RT. The new RTs T'_1 and T'_2 pass the scatterer without interaction, while T'_3 is reflected.

The same algorithm, applied to the RT T_2 , yields two interpolating rays starting at the points $\mathbf{p}_2$ and $\mathbf{p}_3$. We obtain a consistent result if the rays generated within the two RTs, starting from the common point $\mathbf{p}_2$, are the same. Again, as in the point (i) above, our present interpolation ensures identical positions and directions of the new rays, but their curvatures and the associated fields are only approximately equal.

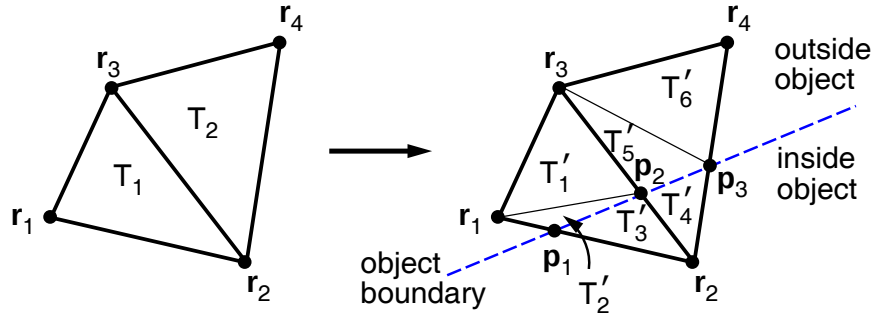


Figure 8: An example of partition of two adjacent RTs due to scattering on the object boundary.

(iv) Even if an incident RT satisfies the angular divergence condition (condition (A) in Sec. 6.2), its divergence may increase as a result of reflection on a curved surface (Fig. 9). We have implemented an algorithm which prevents this situation from occurring. It is illustrated for a simple case in Fig. 10): We first compute the reflected RT corresponding to the original incident RT. We then determine the angles between the reflected rays, and partition the **incident** RT such as to limit the divergence of the reflected RT parts.

Typically, the original RT is partitioned into several smaller RTs defined by the original rays and rays interpolated on the RT edges. As before, the fact that the interpolated rays are located only on common edges of adjacent RTs ensures consistency of RT evolution.

We note that the above algorithm effectively adjusts the transverse sizes of the RTs to the curvature of the scatterer surface at the RT incidence points. If the surface curvature is strongly varying, we can effectively use the initial WF discretization corresponding to the smallest curvature (if

it is sufficient to resolve other geometry details); the considered partition algorithm will then automatically adapt the WF discretization density to the actual variable scatterer curvature.

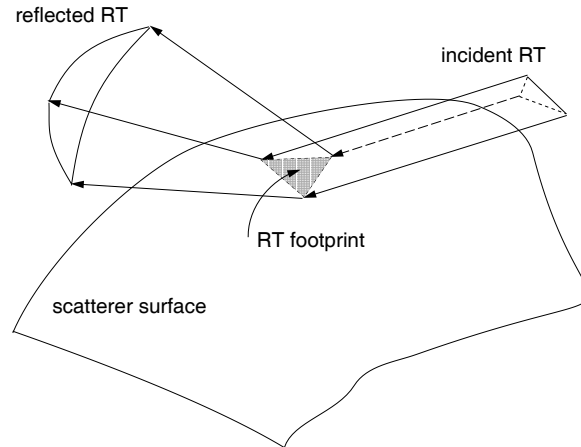


Figure 9: Reflection of a RT on a curved surface, increasing the RT divergence.

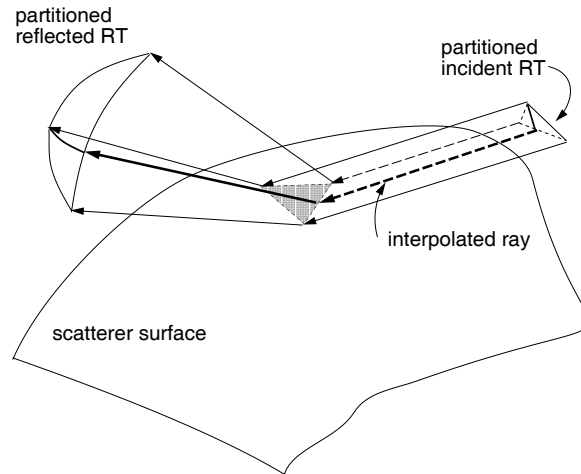


Figure 10: Partition of an incident RT, reducing divergence of the reflected RTs.

(v) Finally, in order to compute the induced currents on the scatterer surface at some prescribed points, we need to construct rays incident exactly at the required surface points. The field associated with these rays are then used to determine the induced currents.

We have implemented an algorithm which, for every RT intersecting the scatterer surface, finds all the required points within the RT footprint, and constructs new rays incident at those points (Fig. 11).

The algorithm involves two elements: (i) a test to determine whether a point is located inside a RT, and (ii) a procedure for constructing a ray emerging from the wavefront of the incident RT, and passing through the given point.

The first element may require nontrivial computations in the common situation when the rays forming the RT tube are not parallel, and the sides of the RT are not flat planes (i.e., when the rays are not coplanar). Our implementation is based on constructing a WF passing near the considered point (using a quadratic interpolation), and then determining positions of the points relative to the planes defined by the boundaries of the WF (Fig. 12). We discuss the procedure of constructing the WF face boundaries in the next Section. Here we only mention that, say, the plane P_{12} of the WF face boundary (curved edge) $(\mathbf{r}_1, \mathbf{r}_2)$ (Fig. 12) is constructed as the least-squares approximation minimizing the angles between the plane and the ray directions $\mathbf{n}_1$ and $\mathbf{n}_2$. The point $\mathbf{r}$ is then considered to be located inside the ray tube if it is located on the interior sides of the three planes, P_{12} , P_{23} , and P_{31} . As follows from this construction, the test is not exact, but, since the divergences of the RTs are, by construction, limited, we found its accuracy entirely sufficient.

The second element of the algorithm will be described in the next Section in the context of ray interpolation methods.

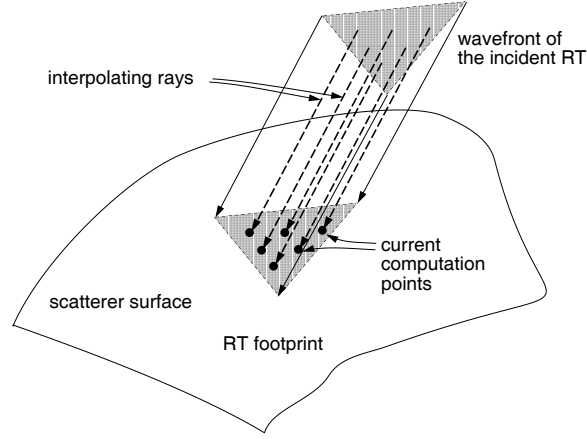


Figure 11: A RT incident on the scatterer surface, with interpolating rays incident at current computation points.

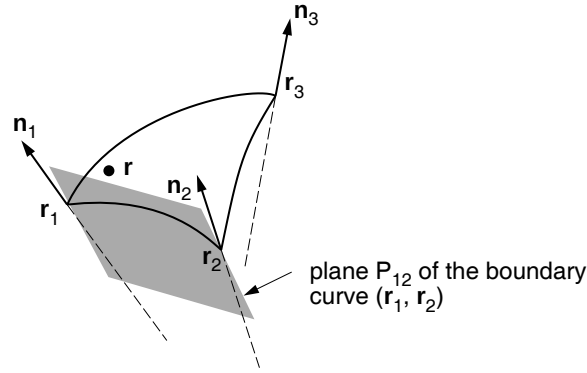


Figure 12: A procedure testing if a point $\mathbf{r}$ is contained inside the given RT.

6.4 Ray tube interpolation algorithms

We implemented two ray interpolation algorithms, mentioned in points (vi) and (vii) in the previous Section. Before we describe these procedures, we comment on our general approach to ray interpolation.

- Since one of the main goals of our WF/RT evolution algorithm is the use of “thick” RTs (not limited in transverse size to a fraction of the wavelength), ray interpolation is an essential element of this approach.

It is required in most of implemented algorithm elements listed in the previous Section.

- The accuracy of ray interpolation should at least correspond to the quadratic representation of the WF inherent in the curvature information associated with the rays. In particular, it is essential to achieve an adequate accuracy of the phases, i.e., the position of the WF.
- Since our computational strategy is to evolve each RT separately, we must ensure consistency of evolution of adjacent RTs, in particular those spanned by interpolated rays.
- In order to comply with the above consistency requirement, we employ the following basic steps in the interpolation procedures:
 1. For a given collection of rays we construct a skeleton of curved edges joining ray points, and defining the WF triangulation.
 2. We specify normals on the edges by interpolating normals at their ends.
 3. We construct face surfaces spanned by edges (boundary curves), using the constructed normals.
 4. Finally, we construct interpolated rays on edges or faces by using, respectively, normals and curvatures defined by the edge curves or the interpolated WF surfaces.

The crucial point of the above algorithm steps is that, since edges are shared by adjacent RTs, the rays interpolated on these edges are constructed consistently.

- In all the procedures mentioned above we try to stay as close as possible to the original local quadratic representation of the WF associated with the ray curvatures. In particular, we take care to minimize higher order (cubic, quartic, ...) effects which could lead to undesirable WF distortions.

We now describe in more detail the interpolation algorithms listed above.

6.4.1 Construction of face boundaries and faces

The first step in our interpolation scheme is the construction of curved edges joining ray vertices $\mathbf{r}_i$. These edges will then form boundary curves of the faces constituting the WF. The procedure is described in Sec. 5.1.

We then use the boundary curve and normal definitions (5.4) and (5.14) to determine the starting points and directions of interpolated rays in one-dimensional interpolation along edges of RT faces. Curvatures of the interpolated rays are then determined on the basis of the interpolated face surfaces, as described in Sec. 5.2.

6.4.2 Interpolation of ray directions

In one-dimensional interpolation of rays on face edges we find new ray directions from the constructed field of normal vectors on the edge (Eq. (5.14)). In two-dimensional interpolation on a WF face we compute ray directions as normals to the surface defined presently by Eq. (5.16).

As we mentioned before, the procedure is not fully consistent, since our present quadratic surface interpolation is not G^1 -continuous. However, discontinuities of normal are numerically small, and will be entirely eliminated by constructing G^1 interpolants as described above.

6.4.3 Interpolation of curvatures

Handling of curvatures in the GO evolution of the WF presents a certain dilemma:

On one hand, the behavior of the fields (typically, their falloff with the propagation distance) is controlled by the **ray curvatures**,

$$\mathbf{E}(s_0 + s) = \sqrt{\frac{\rho_1 \rho_2}{(\rho_1 + s)(\rho_2 + s)}} \mathbf{E}(s_0) , \quad (6.1)$$

where ρ_1 and ρ_2 are the ray's principal curvature radii at the evolution parameter s_0 [5].

On the other hand, the falloff of the fields is controlled, through flux conservation, by the rate of expansion of the RTs, i.e., by their angular divergence. This divergence is related, through our WF surface interpolation procedure described above, to the **WF surface curvature**.

Thus, there may be a potential conflict between using **ray** and **surface** curvatures, if, due to possible approximation errors, the two sets of quantities are not precisely compatible.

In this situation we have decided to adhere to the **surface** curvatures (related to the RT divergence) as more fundamental quantities. Our main argument in favor of this choice is that it ensures consistency of field evolution with the geometrical expansion of RTs. This feature seems to be particularly relevant in the case of reflection off a curved scatterer surface. In that

case reflected ray curvatures are computed on the basis of the scatterer's surface curvatures [5], and the latter may be, in practice, poorly known (especially if the surface is specified as an interpolation of a triangulated surface). Thus, while the ray curvatures may involve a significant error, the geometrical shape (divergence) of the scattered RT is usually better defined and less uncertain.

Therefore, we choose to **discard** the curvature information associated with the rays themselves, and use the RT divergence to construct the interpolated WF, from which eventually the curvatures of rays are determined. This procedure appears to interpolated rays as well.

A disadvantage of this approach is that it does not guarantee continuity of the ray curvatures across the boundaries of the curved triangles (which could have been ensured by interpolating ray curvature data, since rays are common to the adjacent triangles). This does, in general, cause some discontinuities of the surface current; however, the effects are of the second order in h/ρ . Thus, while there is a possibility of further developments, our present judgment is that the advantages of the surface-based approach outweigh its drawbacks, and it is this approach we implemented in our code.

Having a surface parameterization, we compute curvatures in the customary way in terms of the first and second principal forms. For a general parametric surface $\mathbf{r}(u^1, u^2, 1 - u^1 - u^2) = \mathbf{r}(u^1, u^2)$ they are defined by the differentials

$$d\mathbf{r}^2 = g_{\alpha\beta} u^\alpha u^\beta = E du^1 du^1 + 2F du^1 du^2 + G du^2 du^2, \quad (6.2)$$

and

$$-d\mathbf{n} \cdot d\mathbf{r} = L du^1 du^1 + 2M du^1 du^2 + N du^2 du^2, \quad (6.3)$$

where

$$\mathbf{n} = \frac{\partial_1 \mathbf{r} \times \partial_2 \mathbf{r}}{|\partial_1 \mathbf{r} \times \partial_2 \mathbf{r}|} \quad (6.4)$$

is the normal to the surface, $\partial_\alpha \equiv \partial \mathbf{r} / \partial u^\alpha$, and $g_{\alpha\beta}$ is the metric tensor. More explicitly,

$$L = \mathbf{n} \cdot \partial_1^2 \mathbf{r}, \quad M = \mathbf{n} \cdot \partial_1 \partial_2 \mathbf{r}, \quad N = \mathbf{n} \cdot \partial_2^2 \mathbf{r}. \quad (6.5)$$

The two principal curvatures are the two solutions $\kappa = \kappa_1$ and $\kappa = \kappa_2$ of the characteristic equation

$$\det \begin{bmatrix} L - \kappa E & M - \kappa F \\ M - \kappa F & N - \kappa G \end{bmatrix} = 0. \quad (6.6)$$

It follows that the Gaussian curvature (the product of the two principal curvatures) is

$$K \equiv \kappa_1 \kappa_2 = \frac{LN - M^2}{EG - F^2} \quad (6.7)$$

and the mean curvature is

$$H \equiv \frac{1}{2}(\kappa_1 + \kappa_2) = \frac{1}{2} \frac{EN - 2FM + GL}{EG - F^2} . \quad (6.8)$$

In order to obtain the curvatures of the three rays forming the RT, it is sufficient to evaluate the principal curvatures and the curvature directions at one of the vertices, say $\mathbf{r}_3$, at which $u^1 = u^2 = 0$. The remaining expressions can be obtained by permuting the indices.

In our case of the **second-order** surface given by Eq.(5.16) the first derivatives of $\mathbf{r}$ are

$$\partial_1 \mathbf{r} = -\mathbf{R}_2 + \mathbf{b}_1 (1 - 2u^1) + \mathbf{b}_3 [(1 - 2(u^1 + u^2))] , \quad (6.9)$$

$$\partial_2 \mathbf{r} = \mathbf{R}_1 + \mathbf{b}_2 (1 - 2u^2) + \mathbf{b}_3 [(1 - 2(u^1 + u^2))] , \quad (6.10)$$

where

$$\mathbf{R}_1 = \mathbf{r}_2 - \mathbf{r}_3 , \quad \mathbf{R}_2 = \mathbf{r}_3 - \mathbf{r}_1 , \quad \mathbf{R}_3 = \mathbf{r}_1 - \mathbf{r}_2 . \quad (6.11)$$

We can also verify that Eq.(6.4), at $u^1 = u^2 = 0$, yields $\mathbf{n} = \mathbf{n}_3$.

The second fundamental form requires the second derivatives,

$$\partial_1 \partial_1 \mathbf{r} = -2 \mathbf{B}_2 , \quad \partial_1 \partial_2 \mathbf{r} = -2 \mathbf{b}_3 , \quad \partial_2 \partial_2 \mathbf{r} = -2 \mathbf{B}_1 . \quad (6.12)$$

The coefficients of the fundamental forms at $u^1 = u^2 = 0$ are then

$$E = (-\mathbf{R}_2 + \mathbf{B}_2)^2 , \quad F = (-\mathbf{R}_2 + \mathbf{B}_2) \cdot (\mathbf{R}_1 + \mathbf{B}_1) , \quad G = (\mathbf{R}_1 + \mathbf{B}_1)^2 , \quad (6.13)$$

and

$$L = -2 \mathbf{n}_3 \cdot \mathbf{R}_2 , \quad M = -2 \mathbf{n}_3 \cdot \mathbf{b}_3 , \quad N = 2 \mathbf{n}_3 \cdot \mathbf{R}_1 . \quad (6.14)$$

From these coefficients the principal curvatures and curvature directions can be determined using the general formulas, such as Eq.(6.6). Formulae for curvatures of the interpolated rays at arbitrary u^1 and u^2 are only slightly more complicated.

6.4.4 Interpolation of fields

In our representation of fields the rapidly oscillating phase factors are excluded; hence fields on a WF are smooth functions of coordinates that can be reliably interpolated. We currently use a simple linear interpolation of fields $\mathbf{E}_i$ associated with the vertices of a WF face,

$$\mathbf{E}(u^1, u^2, u^3) = \mathbf{E}_1 u^1 + \mathbf{E}_2 u^2 + \mathbf{E}_3 u^3 . \quad (6.15)$$

6.4.5 Computation of surface currents and the scattered field

The interpolated fields are used to compute currents $\mathbf{J}$ (on the surface of a perfect electric conductor), according to the relation $\mathbf{J} = 2 \mathbf{n} \times \mathbf{H}$, with $\mathbf{n}$ constructed as the normal to the surface, and with the magnetic field expressed in terms of the electric field, in our normalization, as $\mathbf{H} = \hat{\mathbf{k}} \times \mathbf{E}$, where $\mathbf{k}$ is the wave vector associated with the ray incident on the surface.

In our present implementation we evaluate currents on a finely discretized scatterer surface (with a discretization ~ 10 points per wavelength). More specifically, using a fine scatterer surface triangulation, we interpolate, within a given RT, a ray incident on a center of each surface face, and evaluate the current at this point. In order to compute the scattered field, we simply sum contributions of all surface faces.

Clearly, this approach can be considered only as an interim solution. It is not acceptable for growing frequencies, since its cost increases (for a single scattering direction) as the frequency squared, incompatibly with the frequency-independent cost of constructing the RTs themselves. In the fully scalable method surface currents should be parameterized as sums of products of smooth and rapidly oscillating functions, and fast methods of integration of these currents, and evaluation of the scattered field should be applied.

6.4.6 Difficulties in ray tube interpolation procedures

We summarize here briefly the difficulties we encountered in constructing WF and RT interpolations. The problems stem from the fact that, in a certain sense, too much information is contained in the WF description in terms of rays:

- (i) there is curvature information associated with the rays themselves; and

- (ii) there is information on the ray directions (i.e., on the RT divergence), from which an interpolating WF surface can be determined, and used in turn to evaluate the curvatures.

Ideally, the information in (i) and (ii) should be consistent, and we should be able to utilize either (i) or (ii) (or both) as input in WF interpolation. In practice, however, there is no guarantee of their compatibility, especially in the presence of various approximations made in the description of RTs and the scatterer.

In our present implementation we opted to discard the information in (i) and use exclusively the information in (ii). Our motivation was that this approach (a) ensures consistency between the rate of decrease in propagating fields and the spreading of the RT, and (b) is less susceptible to uncertainties in determining curvatures of the scatterer surface. A drawback of the selected approach is that it leads, in general, to some discontinuities of the curvatures across boundaries of the adjacent triangles (RTs).

6.5 Numerical examples of ray tube evolution

We present below some numerical results obtained by means of the present version of the WF/RT evolution code.

6.5.1 Ray tube expansion and splitting in free-space evolution

We first consider an example of an expanding RT undergoing partitions which maintain a bounded transverse size of the child RTs. In Fig. 13 we show an initial RT (at evolution parameter $s = 0$) defined as a part of an approximately cylindrical WF with the curvature radii $\rho_1 = 1$ and $\rho_2 = 0$. The Figure shows the intermediate evolution steps leading to the final WF at $s = 5$. The maximum transverse RT size was set here to twice the initial RT size. It is seen that the final WF has a somewhat more complex shape than the initial one, but that the child RTs are adjacent and form a continuous WF surface.

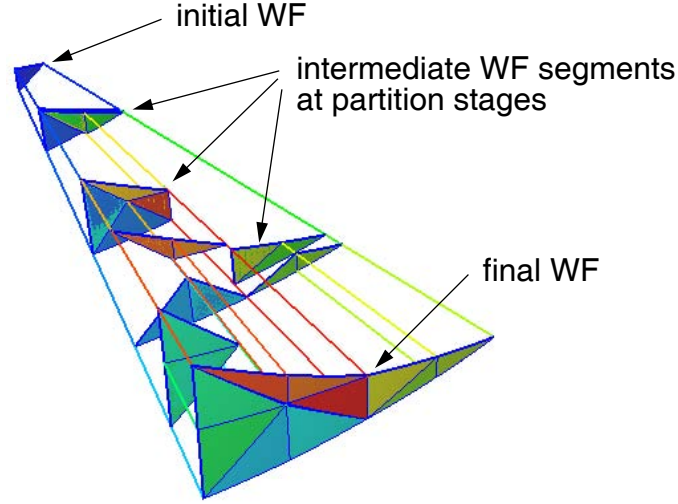


Figure 13: Partitions of an expanding RT evolving in free space.

6.5.2 Ray tube splitting in incidence on object boundary

As an example, we show in Fig. 14 an ellipsoid illuminated by a sample of four adjacent RTs representing a part of a plane-wave WF. The individual RTs are split into reflected and forward-propagating parts in a way illustrated in Fig. 8. Since we did not use here the algorithm for RT partition related to reflection (Figs. 9 and 10), the reflected RTs are highly divergent and span large angles (up to about 60°).

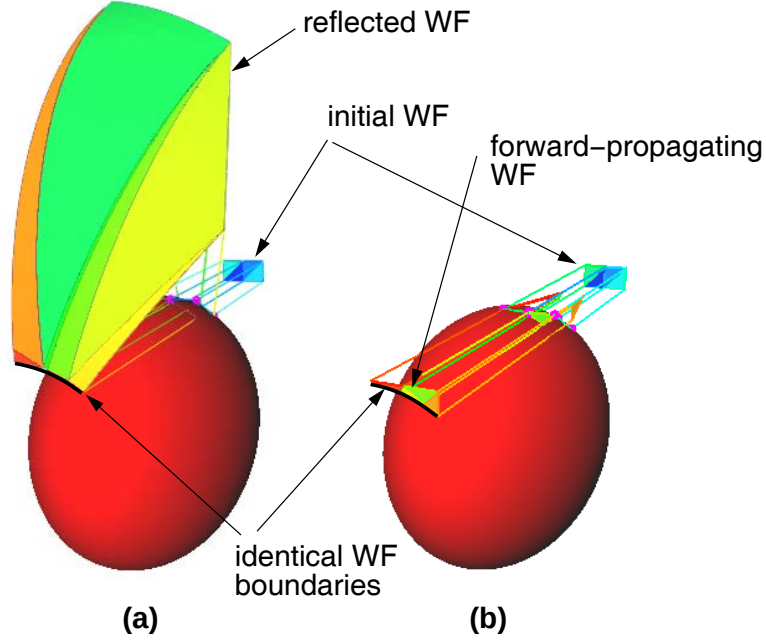


Figure 14: Partitions of RTs reflecting near the object boundary. The initial RTs are split into reflecting parts (a), and forward-propagating parts (b). The black curves mark the common boundary of the reflected and forward-propagating WF segments.

6.5.3 Ray tube splitting in reflections

As a further development of the previous example, we demonstrate now the effect of the algorithm used to partition RT *before* their reflection on a curved surface, in order to limit the divergence of the reflected RTs (the principle of this algorithm is illustrated in Figs. 9 and 10). The result of the algorithm in the considered example is visualized in Fig. 15, obtained by restricting the divergence angles of the reflected RT to 10° . Fig. 15(a) shows the reflected RTs, and Fig. 15(b) the partition of the initial incident RTs (seen in the direction of the RT propagation). It can be seen that the partition of the incident RT tends to refine it close to the projection of the scatterer boundary, i.e., close to near-grazing incidence.

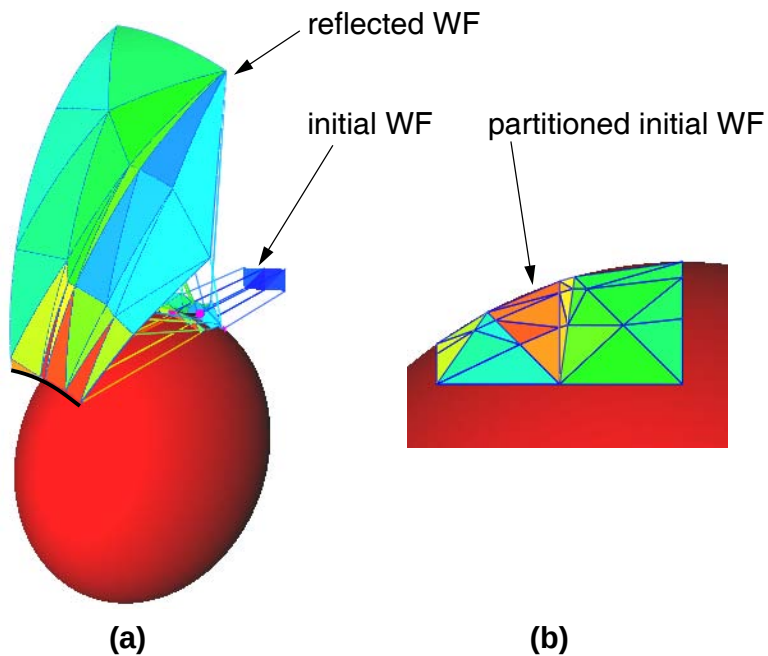


Figure 15: Partitions of RTs related to reflection from a curved surface, for the same configuration as in Fig. 14. The reflected RTs (a) result from the partition (b) of the incident RTs.

6.5.4 Scattering on a system of two smooth convex objects

We give below two examples of computation of scattering cross-section on systems of two smooth convex objects.

The first case, of a system of two ellipsoids, tests the algorithms for ray tube reflection, partition in free-space propagation, partition on the shadow boundary, and adaptive partition prior to reflection (reducing angular divergence of the reflected sub-tubes). In this problem, however, the spreading of the reflected ray tubes is large, and thus the effect of double scattering is small.

In the second case we enhance the double scattering by analyzing a system of two rounded plate-like objects with large flat surface parts, forming a dihedral-type structure. In this case we also compare computations for two frequencies, using *the same* wavefront (ray-tube) discretization.

In both cases we find that, with refining the wavefront discretization, the results practically converge to the limiting cross-section as soon as the ray-

tube width h falls below the value about one-fifth of the minimum curvature radius.

1. A system of two ellipsoids

As the first example of an application of our full WF/RT evolution algorithm we present results of a calculation of scattering cross-section on a system of two ellipsoids, as shown in Fig. 16. The ellipsoids have sizes $2.4 \times 2.0 \times 1.6$ length units, or $8\lambda \times 6.7\lambda \times 3.3\lambda$ for the considered frequency, and are separated by a gap of about 2λ width. The minimum curvature radius is about $\rho_{\min} \simeq 0.533$.

In order to enhance the effect of double reflection, we constructed an incident WF (propagating along the negative z -axis) illuminating only the first ellipsoid, as indicated in Fig. 16. This figure visualizes also a sample ray tube illuminating the first ellipsoid (S_1), reflecting off its surface, and illuminating the second ellipsoid (S_2). The incident and reflected RTs induce surface currents whose y components are shown in the Figure.

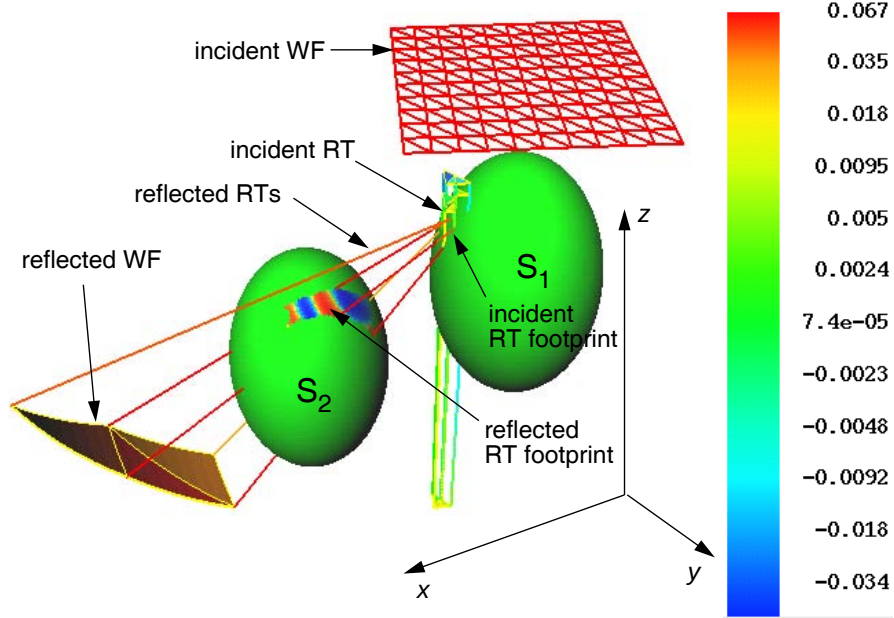


Figure 16: A sample RT (originating from the shown plane-wave WF) illuminating the ellipsoid S_1 , reflecting, and illuminating the ellipsoid S_2 . Shown on the ellipsoids' surfaces is the real part of the current component J_y .

In Figs. 17, 18, and 19 we plot bistatic cross-sections for scattering on the considered system, with the incident plane wave direction along the longest axis of the ellipsoids (as shown in Fig. 16).

Fig. 17 shows the single reflection contribution computed for two resolutions h of the WF discretization (average ray-ray distance): $h = 0.2$ and $h = 0.1$. Similarly, Fig. 18 shows the single and double reflection contributions computed for WF resolutions $h = 0.2$, $h = 0.1$, and $h = 0.05$ (corresponding to $0.375 \rho_{\min}$, $0.187 \rho_{\min}$, and $0.094 \rho_{\min}$).

In Fig. 19 we indicate the effects of double reflection by comparing the cross sections for single and single+double reflection, both computed for $h = 0.1$. For the considered geometry and the incident wave there are, actually, no contributions of triple and higher order reflections.

Fig. 19 shows a definite, although not large, influence of double reflection. It appears that its effect is mainly to reduce the scattering cross section in the angular range in which the rays reflected from the ellipsoid S_1 are shadowed by the ellipsoid S_2 .

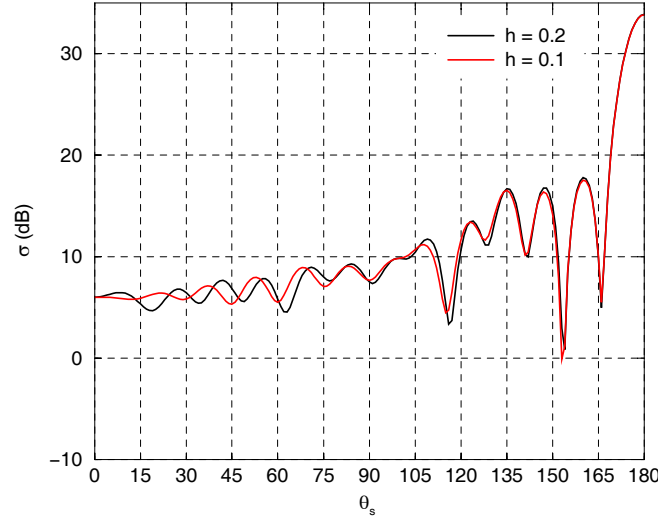


Figure 17: Single reflection contribution to the bistatic scattering cross-section on the system of two ellipsoids, for two different WF discretizations.

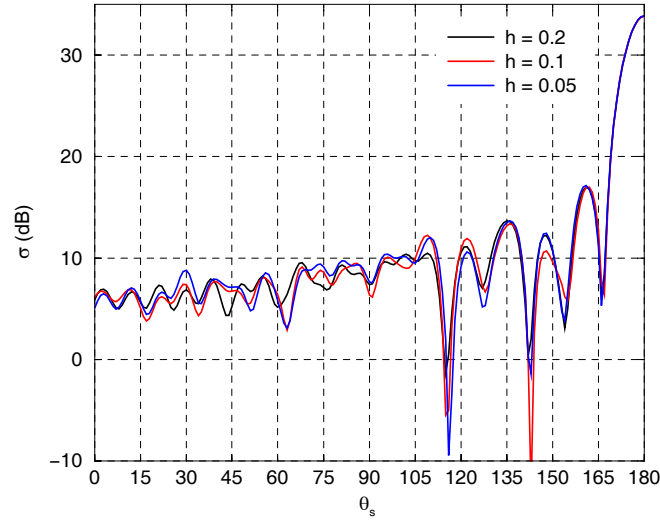


Figure 18: Single+double reflection contribution to the bistatic scattering cross-section on the system of two ellipsoids, for three different WF discretizations.

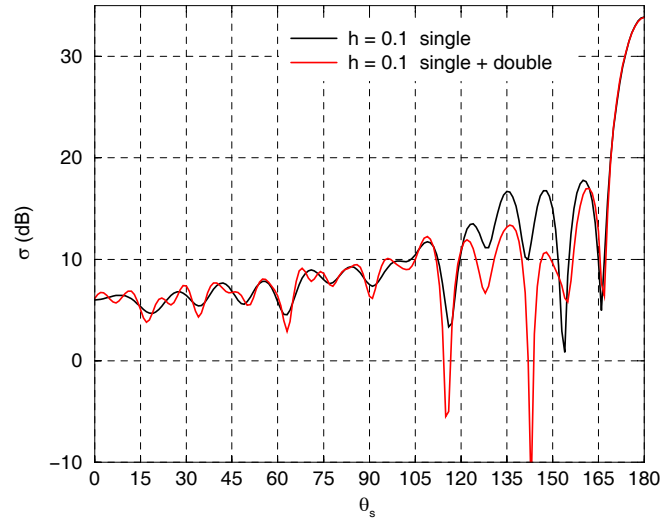


Figure 19: Comparison of single and single+double reflection contributions to the bistatic scattering cross-section on the system of two ellipsoids, computed for the same WF discretization.

We note that, although the convergence of the cross-section results with the WF resolution is not fully satisfactory, the discrepancies occur mainly for near-back-scattering directions ($\theta_s \simeq 0$), where single scattering dominates. In the region of angles $110^\circ \lesssim \theta_s \lesssim 160^\circ$ where double reflections are important (Fig. 19) the results appear to weakly dependent on the WF discretization resolution h (Fig. 18), as soon as $h \lesssim \rho_{\min}/5$.

We also note that the double reflection contribution is being tested in a rather difficult case, when the primary incident RTs reflect from the first ellipsoid at almost grazing direction (Fig. 16), and are often incident on the second one also at near-grazing angles. One of the consequences of this geometrical configuration are strongly elongated shapes of the RT footprints of the RTs, as also shown in Fig. 16. Such footprint shapes constitute a rather stringent test of our field interpolation procedure.

2. A system of two rounded plates

In the previously considered example of scattering on a system of two ellipsoids the effects of double reflection, although visible, were not strong. The reason was that both single and double reflection were generating divergent ray tubes, in which the scattered field was rather rapidly decaying with the propagation distance.

Therefore, in order to enhance the double reflection effect, we considered another case of scattering on a system of two flat thick “plates” with rounded edges and corners (which makes our present smooth-surface algorithm applicable), forming a dihedral-type structure. The scatterer, together with a sample ray tube, is shown in Fig. 20. The two plates are located in the (x, z) and (y, z) planes, and are illuminated with a plane wave incident symmetrically at the angles $\theta_i = 90^\circ$ and $\phi_i = 45^\circ$, with the electric field along the z axis (i.e., the vertical polarization). The sizes of the plates are, in some length units, $7 \times 7 \times 1$, and the minimum curvature radius is $\rho_{\min} = 0.5$.

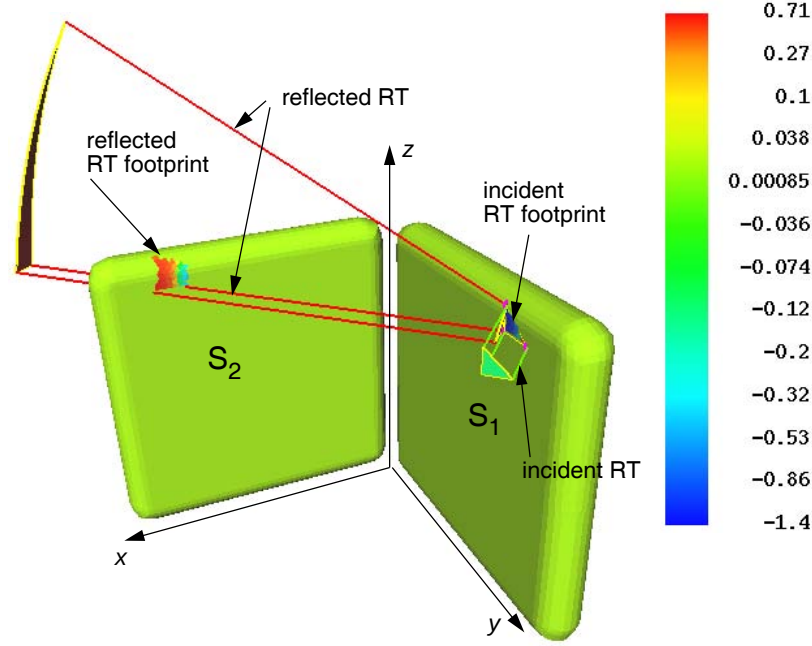


Figure 20: A sample RT illuminating the plate S_1 , reflecting, and illuminating the plate S_2 , for the incident WF resolution $h = 0.5$. Shown on the surfaces is the imaginary part of the current component J_z .

We carried out the first computation at the wavelength $\lambda = 1.5$. Fig. 21 shows the bistatic cross-section, as a function of ϕ_s and at $\theta_s = 90^\circ$. The contributions of single and single+double reflections are shown in for three WF resolutions (ray-ray distances): $h = 1.0$, $h = 0.5$, and $h = 0.25$ (i.e., $2.0 \rho_{\min}$, $1.0 \rho_{\min}$, and $0.5 \rho_{\min}$). While the coarsest resolution ($h = 2/3 \lambda$) does not provide an adequate accuracy of the double reflection contribution, the difference between the remaining finer discretizations is quite small (less than 1 dB).

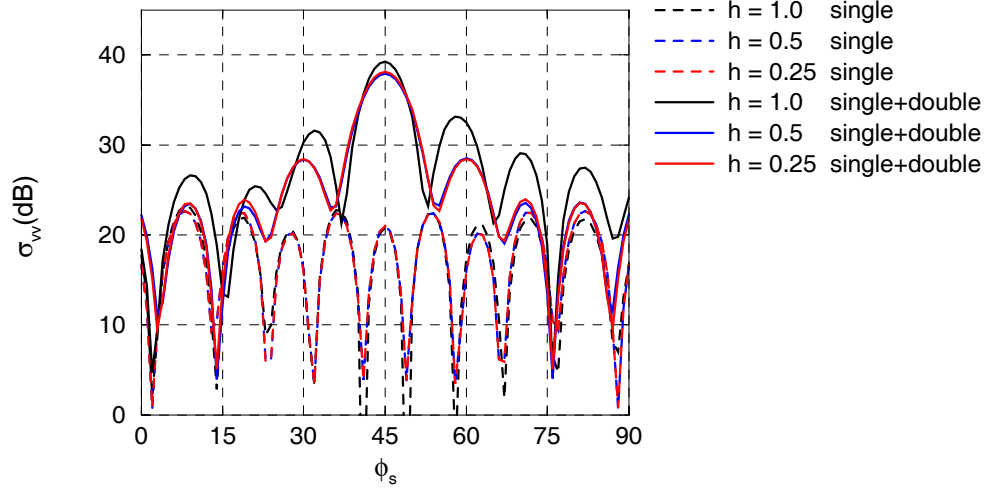


Figure 21: Comparison of single and single+double reflection contributions to the bistatic scattering cross-section on the system of two flat plates, computed for three WF discretizations.

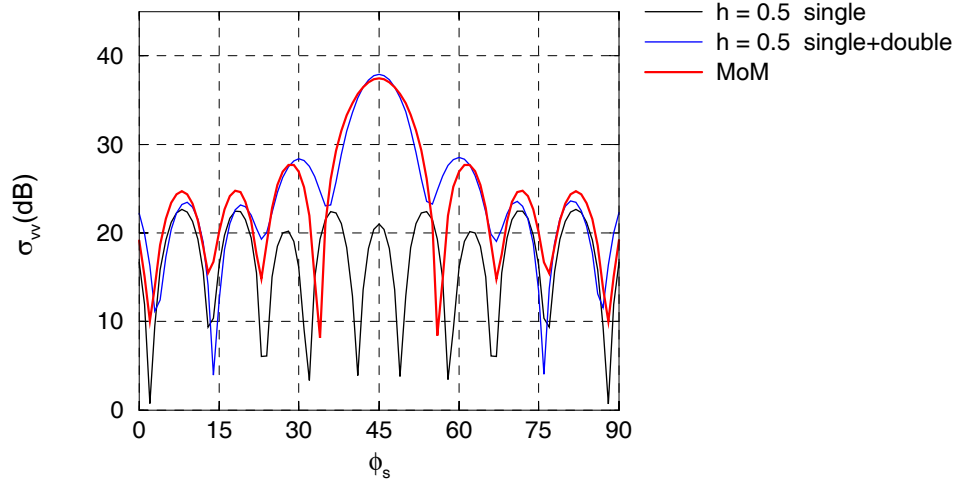


Figure 22: Single and single+double reflection contributions, computed for the intermediate resolution ($h = 0.5$), compared with the rigorous MoM result.

Fig. 21 clearly indicates a strong effect of double reflection which, as expected, causes a significant enhancement of the cross-section in the back-scattering direction. In Fig. 22 we compare the single and single+double reflection results with the rigorous MoM computation (obtained with about 43,000 unknowns). The comparison shows that the contributions other than reflections (i.e., several diffraction mechanism) are relatively less important.

We verify now, for the considered example, that the WF discretization, once adjusted to the given scatterer geometry, can be used ***independently of the frequency***. To this end, we used the same WF discretization as that used to obtain the results of Figs. 21 and 22, and increased the frequency by the factor of two (corresponding to the wavelength $\lambda = 0.75$).

In Fig. 23 we show the resulting single and single+double reflection contributions to the bistatic cross-section, compared with the rigorous MoM result (obtained with about 170,000 unknowns). The accuracy of the RT computation appears to be about the same as at the lower frequency; the differences in the results for the WF resolutions $h = 0.5$ and $h = 0.25$ are, again, less than 1 dB.

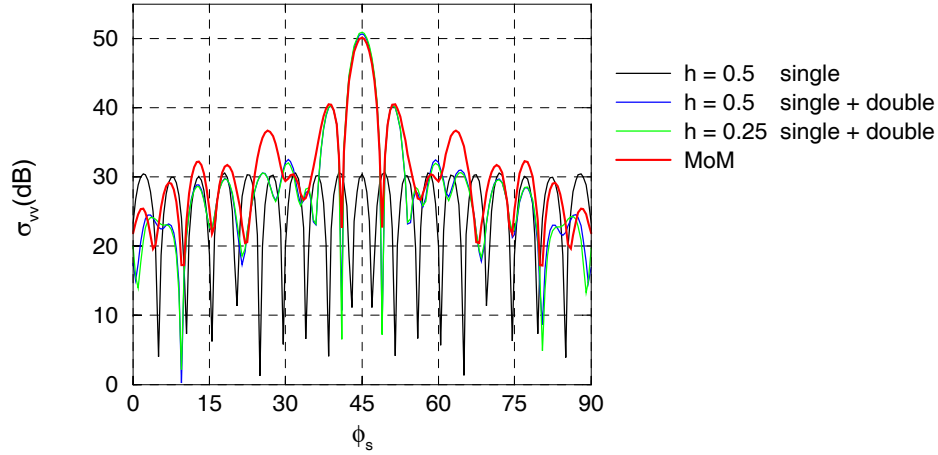


Figure 23: Single and single+double reflection contributions (the latter computed for two resolutions, $h = 0.5$ and $h = 0.25$), compared with the rigorous MoM result, computed at twice the frequency of Figs. 21 and 22.

7 Technical details: C. Evolution of creeping ray tubes (ray-strips)

We have developed and implemented a set of algorithms which allow us to accurately trace creeping rays and ray strips on smooth convex surfaces, and to evaluate surface currents associated with these rays.

7.1 Curved surface parameterization

In our algorithms ray-tube evolution, we consistently utilize second-order parameterization and interpolation of the wavefront segments. In this context it was natural to select a compatible representation of the scatterer surface, which would lead to manageable analytic representation of rays and ray tubes. We use thus the second-order parameterization described in Sec. 5.

7.2 Construction of creeping rays (geodesics)

Our algorithm for generating creeping rays has been devised for closed convex surfaces parameterized by means of second-order curved triangles.

In principle, the creeping rays (geodesics of the surface) can be generated by solving the geodesic equation. If a geodesic, as a curve in the parameter space $\mathbf{u} = (u^1, u^2)$, is denoted by

$$\mathbf{u}(\sigma) \equiv (u^1(\sigma), u^2(\sigma)) , \quad (7.1)$$

it can be obtained as a solution of the second-order vector nonlinear differential equation

$$\frac{d^2 u^\alpha}{d\sigma^2} + \Gamma^\alpha_{\beta\gamma} \frac{du^\beta}{d\sigma} \frac{du^\gamma}{d\sigma} = 0 , \quad (7.2)$$

where the summation over repeated indices is understood, and the Christoffel symbols (symmetric in the two lower indices) are given, in terms of the metric tensor g , by

$$\Gamma^\alpha_{\beta\gamma} \equiv g^{\alpha\delta} \Gamma_{\delta\beta\gamma} = g^{\alpha\delta} \frac{1}{2} (\partial_\beta g_{\delta\gamma} + \partial_\gamma g_{\delta\beta} - \partial_\delta g_{\beta\gamma}) , \quad (7.3)$$

or, in two dimensions, in terms of the coefficients of the fundamental forms

(Eqs. 6.13 and 6.14),

$$\Gamma^1_{11} = \frac{G \partial_1 E - 2F \partial_1 F + F \partial_2 E}{2(EG - F^2)}, \quad (7.4)$$

$$\Gamma^2_{11} = -\frac{F \partial_1 E - 2E \partial_1 F + E \partial_2 E}{2(EG - F^2)}, \quad (7.5)$$

$$\Gamma^1_{12} = \frac{G \partial_2 E - F \partial_1 G}{2(EG - F^2)}, \quad (7.6)$$

$$\Gamma^2_{12} = \frac{E \partial_1 G - F \partial_2 E}{2(EG - F^2)}, \quad (7.7)$$

$$\Gamma^1_{22} = -\frac{F \partial_2 G - 2G \partial_2 F + G \partial_1 G}{2(EG - F^2)}, \quad (7.8)$$

$$\Gamma^2_{22} = \frac{E \partial_2 G - 2F \partial_2 F + F \partial_1 G}{2(EG - F^2)}. \quad (7.9)$$

Equations (7.2) determine a unique geodesic passing through any specified point in any given direction.

With the second-order surface parameterization the elements of the Christoffel symbol are rational functions of the parameters $\mathbf{u}$. One possible approach to solving the geodesic equations (7.2) is to expand the coefficients and the solution as polynomials in the evolution parameter σ and the coordinates $\mathbf{u}$. In the implementation of this approach, however, we found that, even for moderately curved triangles, the coefficients in the geodesic equation, although smooth and regular, exhibit a rather strong variation within the individual curved triangles. This circumstance requires high expansion orders (and complicated expressions for the expansion coefficients), and renders the analytical solution impractical. In this situation we have to resort to a numerical solution. However, instead of solving the geodesic equations numerically in the parameter space, we have chosen an equivalent (and in practice simpler) procedure of performing the ray tracing directly on the finely discretized curved triangle surface, and mapping the obtained spatial geodesic back to the parameter space. The procedure amounts to the following steps:

1. We construct on the unit triangle $\mathcal{T}$ in the parameter space a regular $K \times K$ structured grid of smaller sub-triangles.
2. We map vertices of the sub-triangles in the $\mathbf{u}$ -space onto the curved triangle, and construct in this way an approximation of the curved triangle surface T as a set of small triangular facets.

3. Considering the facets of the triangle T as flat, we perform the conventional ray tracing, in which rays consist of straight-line segments on the individual facets. Upon crossing the facet boundaries, the rays satisfy the condition (following from the Fermat principle) of equal angles between the edge and the rays. The same conditions are applied to rays crossing boundaries between the individual curved triangles.
4. Finally, we find a representation of the obtained ray (geodesic) trajectory in the parameter space $\mathbf{u}$ by using the mapping $\mathbf{r} \longrightarrow \mathbf{u}$ inverse to that of Eq. (5.16). For efficiency of further manipulations, we approximate the numerically determined trajectory as a quadratic function of the evolution parameter σ .

We found that the geodesic construction described above incurs low computational costs and provides high accuracy already for K between 10 and 20, for curved triangles of sizes comparable to their curvature radii.

An illustration of the accuracy of the implemented method we show in Fig. 24 a sample set of creeping rays (geodesics) arriving at one tip of the ogive (Fig. 3), after being launched from the opposite tip (physically, due to tip diffraction). As in Figs. 3 and 4, the ogive surface was interpolated with 800 curved triangles of edge lengths equal to about 1/25-th of the ogive length.

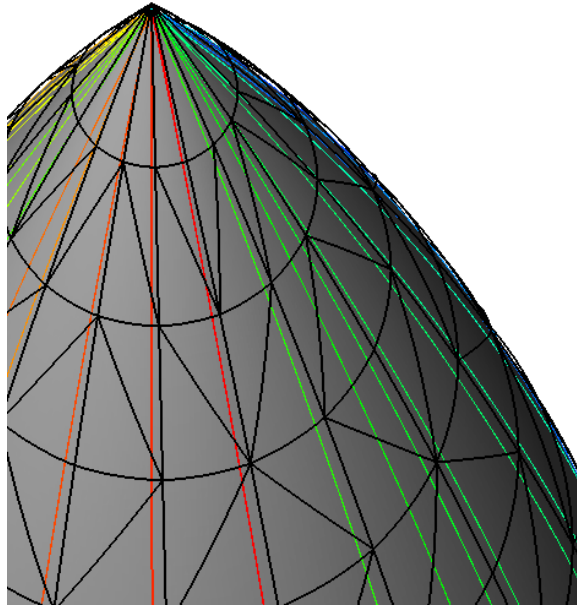


Figure 24: Behavior (in the vicinity of an ogive's tip) of a sample of creeping rays (geodesics) emerging from the other tip.

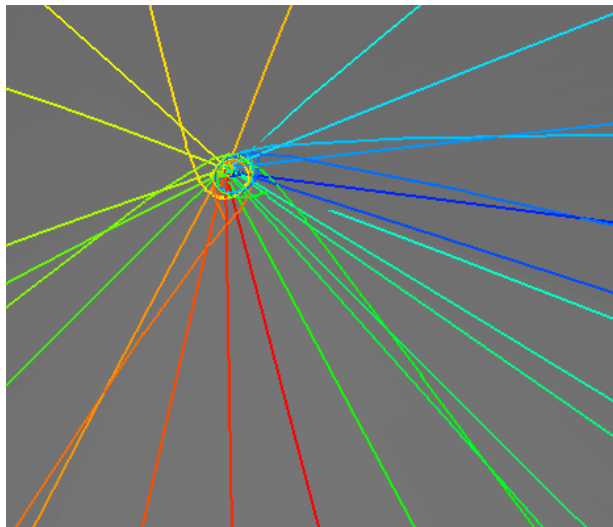


Figure 25: An enlargement (by approximately a factor 10) of the vicinity of the ogive tip, with the set of geodesics shown in Fig. 25.

At the resolution of Fig. 24 the errors in the geodesic construction are visually undetectable. An approximately ten-fold enlargement shown in Fig. 25 reveals that some of the geodesics miss the tip and form loops around it. However, the radii of these loops do not exceed about 5 mm (compared to the 10 m ogive length).

7.3 Evolution of ray strips

An important element of the creeping ray (geodesics) construction is to determine the rate at which the nearby rays diverge or converge when traveling on the surface; this rate controls the changes of magnitude of the surface-diffracted fields associated with the rays.

The behavior of the nearby creeping rays corresponds to the behavior of ray tubes, which, on the surface, degenerate to “ray strips”. It can be conveniently described for an infinitesimally narrow ray strip, i.e., a pair of infinitesimally displaced geodesics. Such a pair is characterized by the well-known *geodesic-deviation equation* – a second-order linear equation for the separation between the geodesics, involving the Riemann-Christoffel curvature tensor. In the considered case the equation for the ray strip (infinitesimal) width $w(s)$ involves the Gaussian curvature K of the surface. For $K > 0$ the width of the ray strip, as a function of the evolution parameter

s , has an oscillatory behavior, while for $K < 0$ is exponentially growing or decreasing (clearly, $K \geq 0$ for a convex surface).

More specifically, if the geodesic curve $\mathbf{u}(s)$ is parameterized by its length s , the width of the strip satisfies the equation

$$\ddot{w}(s) + K(\mathbf{u}(s)) w(s) = 0 \quad \text{for } \sigma = s . \quad (7.10)$$

If we interpret s as time, this equation is (for a non-negative Gaussian curvature, which is the for a convex surface) the harmonic oscillator equation with a time-dependent “spring constant” K . In this case the solutions have an oscillatory character and describe neighboring geodesics periodically converging at focal points (for a negative Gaussian curvature the solutions would have an exponentially growing or decreasing character).

In our case of a quadratically parameterized surface we expand the solution to Eq. (7.10) to the fourth powers in the evolution parameter σ and use the corresponding expansion for the curvature, which can be obtained from the quadratic expansion of the geodesic. We similarly expand the geodesic speed squared $(ds/d\sigma) = v_0 + v_1 \sigma + \dots$, where σ is defined relative to reference point within the considered triangle. To this order, the solution for the geodesic deviation is

$$w(\sigma) = w(0) \left(1 - \frac{1}{2} K_0 v_0 \sigma^2\right) + \dot{w}(0) \sigma + \dots , \quad (7.11)$$

where K_0 is the Gaussian curvature at the reference point $\sigma = 0$.

In addition to solving the geodesic deviation equation on individual triangles, we also impose exact continuity conditions at junctions of the triangles. This procedure requires some care, since, at edges at which the surface is not G^1 -continuous, the Gaussian curvature has a delta-function singularity resulting in a discontinuity of $dw(s)/ds$.

7.4 Implementation of the creeping-ray construction algorithms

As we mentioned, the algorithms for creeping-ray construction result in local analytic (second-order) parameterizations of the rays and ray strips on the individual curved triangles. A ray is specified in the surface parameter space as $\mathbf{u}(s)$, and the ray-strip width as $w(s)$. The parameters defining these functions are stored for all the curved triangles, to be used in further computations – such as surface current evaluation. The implemented ray construction procedure is, briefly, as follows:

1. Determine the shadow boundary as a sequence of edges separating illuminated and shadowed faces (for this purpose the triangles are assumed flat).
2. Locate “sharp”, i.e., “diffraction” edges of the surface.
3. Locate corners and tips of the surface.
4. From the center of each edge on the shadow boundary launch a single surface ray (if the edge is *not* a diffraction edge), or two rays (if the edge is a diffraction edge).
5. Evolve each ray, keeping track of its approximate Fock parameter and the number of “forward” diffraction and “backward” diffraction events. These diffraction events occur when the ray intersects a diffraction edge: then the original ray continues traveling in the “forward” direction, and a new ray is created, traveling in the “backward” direction. The new ray inherits the history of the parent ray (the Fock parameter and the numbers of edge diffractions).
6. Each ray is evolved until it reaches the prescribed bound on the Fock parameter or on the number the edge-diffraction events. In the evolution process the differential equation for the corresponding strip width is solved.
7. Each triangle keeps a register of properties (directions, polarizations, curvatures) of rays that have traversed that triangle. Width parameters of the corresponding ray strips are also stored.
8. After building the first generation of rays, the second generation is constructed by doubling the number of launched rays. This procedure is being repeated until the numbers of “different” rays on faces stabilize.

As an illustration of the ray-strip construction procedure described above we present some ray (geodesic) tracing results for an ellipsoid similar to that considered previously (Fig. 2). The sizes of the ellipsoid are $9 : 6 : 4$, and it is illuminated by a plane wave incident along its longest axis. Fig. 26 shows two ray strips (visualized as a finite width set of three rays), emerging from two nearby points on the shadow boundary. Another view of the set of ray strips is shown in Fig. 27. The evolution of the ray strips was computed using the algorithm described above, and the ellipsoid surface was interpolated using 456 curved triangles.

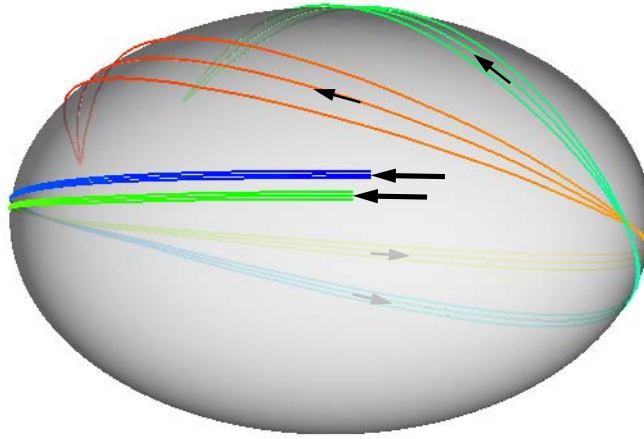


Figure 26: Visualization of two nearby ray strips propagating on the surface of an ellipsoid. The heavy arrows indicate the direction of the incident plane wave.

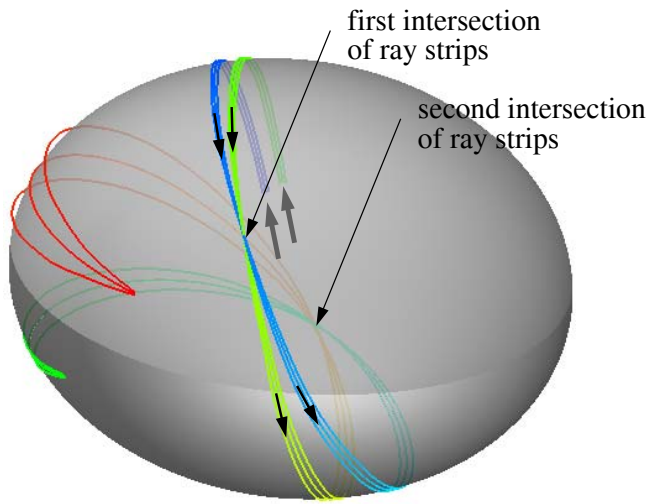


Figure 27: Another view of the set of ray strips of Fig. 26, with indicated intersections of rays.

The Figures show that the analytically evaluated widths of the ray strips (obtained by solving the geodesic deviation equation) are in close correspon-

dence with the numerically determined evolution of neighboring rays. In particular, the first intersection of rays practically coincides with the points at which the widths of the ray strips vanish. At the second intersection there is a somewhat larger discrepancy between the ray intersection points and the point where the ray strips shrink to zero width. The points where the ray strips shrink to zero for the third time are quite widely spaced. However, that happens only after the ray strips had traveled about $1\frac{1}{2}$ times around the object; and, in any case, we cannot expect an exact agreement between the behavior of an infinitesimally narrow ray strip, and a set of rays separated by a finite distance.

7.5 Algorithms for current parameterization and computation

The algorithms described in Secs. 7.1 to 7.4 result in a set of creeping rays and the associated ray strips for the dominant diffraction processes that may occur on a convex perfectly conducting object: smooth-surface, edge, corner, and tip diffraction mechanisms (less important mechanisms, such as diffraction on a surface curvature discontinuity, are not included). In the current implementation we included only smooth surface and edge diffraction, and in the following description we concentrate on computation of surface fields induced by *creeping rays due to smooth surface diffraction* only.

7.6 Asymptotic expressions for the fields and currents due to creeping rays

We consider fields and currents at the point $\mathbf{r} = \mathbf{r}(s)$ on the geodesic, corresponding to the geodesic evolution parameter s . It is convenient to utilize the local coordinate system associated with the geodesic, defined by the triplet of unit vectors $(\mathbf{t}, \mathbf{n}, \mathbf{b})$, where $\mathbf{n}$ is normal to the surface, the $\mathbf{t}$ tangential to the ray, and the binormal vector $\mathbf{b}$ is defined as $\mathbf{b} = \mathbf{t} \times \mathbf{n}$.

We denote by $\mathbf{r}_0$ the point on the ray located at the shadow boundary, corresponding to the evolution (ray length) parameter s_0 .

It is customary to express the fields associated with the creeping ray in terms of their “soft” (s) and “hard” (h) components, corresponding to the transverse-magnetic (TM), or E-polarization, and transverse-electric (TE), or E-polarization, problems in the canonical problem of the perfectly conducting cylinder. Alternatively, the soft component is the field component along the binormal vector $\mathbf{b}$ and the hard component is the component along the normal vector $\mathbf{n}$ (the ray-optical field is, by definition, normal to the ray

propagation direction $\mathbf{t}$).

The well known asymptotic expressions for induced currents associated with creeping waves in the shadow region of perfectly conducting current have been originally derived by Fock in two- and three-dimensional problems [7] (see also Ref. [8] for a higher order expansion). We follow here conventions of Ref. [9], in which the expressions for the creeping-wave currents are derived based on the analysis of the two-dimensional cylinder problem.

The result can be summarized as follows:

The surface electric current $\mathbf{J}$ is related to the magnetic field by $\mathbf{J} = -2\mathbf{n} \times \mathbf{H}$, where $\mathbf{n}$ is the outer normal to the surface. Equivalently, the transverse current component is expressible in terms of the normal derivative of the soft (tangential) electric field component, and the longitudinal current component in terms of the hard (normal) electric field component. The total current is then

$$\mathbf{J}(\mathbf{r}) = \mathcal{T}(\mathbf{r}, \mathbf{r}_0) \cdot \mathbf{E}^i(\mathbf{r}_0) , \quad (7.12)$$

where $\mathbf{E}^i(\mathbf{r}_0)$ is the incident electric field at the shadow boundary point $\mathbf{r}_0$ and $\mathcal{T}(\mathbf{r}, \mathbf{r}_0)$ is dyadic “field-to-current” transfer function of the form

$$\begin{aligned} \mathcal{T}(\mathbf{r}, \mathbf{r}_0) &= [\mathbf{b} \mathbf{b}_0 \mathcal{T}_s(\mathbf{r}, \mathbf{r}_0) + \mathbf{t} \mathbf{n}_0 \mathcal{T}_h(\mathbf{r}, \mathbf{r}_0)] \left[\frac{\rho_g(s_0)}{\rho_g(s)} \right]^{1/6} \left[\frac{w(s_0)}{w(s)} \right]^{1/2} e^{ik(s-s_0)} \\ &\equiv [\mathbf{b} \mathbf{b}_0 \mathcal{T}_s(\mathbf{r}, \mathbf{r}_0) + \mathbf{t} \mathbf{n}_0 \mathcal{T}_h(\mathbf{r}, \mathbf{r}_0)] Q(s, s_0) , \end{aligned} \quad (7.13)$$

where $w(s_0)$ and $w(s)$ are the ray strip widths at the shadow boundary point $\mathbf{r}_0$ and at the field observation point $\mathbf{r}$, and where ρ_g is the radius of the surface curvature measured in the direction of the geodesic, i.e., in the plane $(\mathbf{n}, \mathbf{t})$. In terms of the principal curvatures radii ρ_1 and ρ_2 of the surface at that point, we have

$$\frac{1}{\rho_g} = \frac{\cos^2 \alpha}{\rho_1} + \frac{\sin^2 \alpha}{\rho_2} , \quad (7.14)$$

where α is the angle between the geodesic and the first main curvature direction of the surface. Further, the functions $\mathcal{T}_s$ and $\mathcal{T}_h$ specify the field evolution for the soft and hard components,

$$\mathcal{T}_s(\mathbf{r}, \mathbf{r}_0) = f(\xi(s - s_0)) \left[\frac{1}{2} k \rho_g(s) \right]^{-1/3} \quad (7.15)$$

and

$$\mathcal{T}_h(\mathbf{r}, \mathbf{r}_0) = -ig(\xi(s - s_0)) , \quad (7.16)$$

where the dimensionless *shadow Fock parameter* ξ is

$$\xi(s - s_0) = (\tfrac{1}{2}k)^{1/3} \int_{s_0}^s \frac{ds'}{[\rho_g(s')]^{2/3}} , \quad (7.17)$$

and the functions f and g , referred to as the soft and hard Fock functions [10], are defined as

$$f(\xi) = \frac{1}{\sqrt{\pi}} \int_{\Gamma} du \frac{e^{i\xi u}}{w_1(u)} , \quad (7.18)$$

and

$$g(\xi) = \frac{1}{\sqrt{\pi}} \int_{\Gamma} du \frac{e^{i\xi u}}{w'_1(u)} \quad (7.19)$$

in terms of the Fock-type Airy function

$$w_1(u) = \frac{1}{\sqrt{\pi}} \int_{\Gamma} dt e^{ut - t^3/3} . \quad (7.20)$$

In the equations above the contour Γ consists of two segments: from $\exp(i2\pi/3)\infty$ to 0, and from 0 to ∞ . The asymptotic behavior of the functions f and g for $\xi \rightarrow \infty$ is given by

$$f(\xi) \propto e^{-i\pi/3} \frac{1}{\text{Ai}'(-\alpha_1)} e^{\alpha_1 \xi \exp(5i\pi/6)} \quad (7.21)$$

and

$$g(\xi) \propto \frac{1}{\beta_1 \text{Ai}(-\beta_1)} e^{\beta_1 \xi \exp(5i\pi/6)} , \quad (7.22)$$

where $\alpha_1 \simeq 2.338$ is the negative of the first zero of the Airy function Ai , and $\beta_1 \simeq 1.019$ is the negative of the first zero of the derivative Ai' of that function. The relevant values of the Airy function and its derivative are $\text{Ai}(-\beta_1) \simeq 0.536$ and $\text{Ai}'(-\alpha_1) \simeq 0.701$.

The values of the Fock functions at zero argument are

$$f(0) \simeq 0.776 e^{-i\pi/3} , \quad g(0) \simeq 1.399 . \quad (7.23)$$

Physically, the current (7.12) is interpreted as the “radiating current”, i.e., the current responsible for radiation of the rays “shed” by the ray strip

propagating on a curved surface. Therefore, the current vanishes when the surface curvature $1/\rho_g(s)$ at the radiation point vanishes. That behavior is different for the soft and hard components: according to Eqs. (7.13), (7.15), and (7.16),

$$\mathbf{b} \cdot \mathbf{J} \sim \rho_g^{-1/2} \quad (\text{soft}) \quad (7.24)$$

and

$$\mathbf{t} \cdot \mathbf{J} \sim \rho_g^{-1/6} \quad (\text{hard}) , \quad (7.25)$$

i.e., the radiation strength for the hard component is much less sensitive to the curvature than for the soft component.

It follows from Eqs. (7.21) and (7.22) that the soft component of the current decays in the shadow region more rapidly than the hard component ($\alpha_1 \gtrsim 2\beta_1$). We note that, since $\text{Im} \exp(5i\pi/6) > 0$, the functions $f(\xi)$ and $g(\xi)$ represent not only a damping of the creeping wave, but also provide an additional phase variation: effectively, the propagating wave oscillates faster than it would follow from the phase factor $\exp(iks)$ associated with the optical path. Again, that effect is stronger for the hard component.

We also note that Eq.(7.12) predicts an infinite induced current at the ray-strip caustics (or focal points), where $w(s) \rightarrow 0$. Strictly speaking, if we compute the radiated field by integrating the current, this singularity will be integrable, since, near a caustic singularity $w(s_c) = 0$, the surface integral of the current is proportional to

$$\int_0^{\infty} dt \int_0^{w(s_c+t)} db \frac{1}{\sqrt{w(s_c+t)}} = \int_0^{\infty} dt \sqrt{w(s_c+t)} , \quad (7.26)$$

where $t = s - s_c$ and b denotes the local coordinate perpendicular to the ray direction.

7.7 Parameterization of the current

In the context of our algorithm implementation it is convenient to parameterize the currents in terms of functions defined on the individual curved triangles T (labeled in the following by indices τ). As the *simplest parameterization*, we assume that the current on the triangle T_τ is given by

$$\mathbf{J}_\tau(\mathbf{r}) = \chi_\tau(\mathbf{r}) \sum_{m=1}^M \mathbf{p}_{\tau m} e^{i\mathbf{q}_{\tau m} \cdot (\mathbf{r} - \mathbf{r}_{0\tau})} \equiv \chi_\tau(\mathbf{r}) \sum_{m=1}^M \mathbf{p}_{\tau m} \mathbf{J}_\tau^{(m)}(\mathbf{r}) , \quad (7.27)$$

where $\mathbf{r}_{0\tau}$ is a reference point on the triangle T_τ , χ_τ is the triangle's characteristic function, m labels contributions of various ray strips traversing the triangle, and corresponding to various physical high-frequency mechanisms. Each contribution is characterized by the “polarization vector” $\mathbf{p}$, approximately tangential to the surface, and by the wave vector $\mathbf{q}$ (which includes an imaginary part, associated with the extinction rate of the ray on the given triangle).

We can now express the parameters $\mathbf{p}_{\tau m}$ and $\mathbf{q}_{\tau m}$ in terms of the quantities associated with the m -th ray strip traversing the triangle T_τ . In the following we discuss a contribution of one representative ray; the total current on the triangle is the sum of such contributions.

The relevant parameters of each ray strip are:

1. The value s_τ of the evolution parameter s (optical path length of the ray), relative to some reference wavefront. For a plane incident wave we set $s = 0$ on the wavefront passing through the coordinate system origin. The parameter s associated with a creeping ray refers to the starting point, say $\mathbf{r}_\tau$, of the ray on the triangle (the point at which the ray enters the triangle surface). The evolution parameter s is being computed numerically in the ray tracing procedure by summing increments of the geodesic length (using the refined discretization of the curved triangles).
2. The value ξ_τ (at the point $\mathbf{r}_\tau$) of the Fock parameter accumulated by the ray. This parameter is also being computed in the ray tracing by accumulating contributions of triangles traversed by the ray, according to Eq.7.17, using average triangle curvatures. The result is well defined, even though in our surface parameterization the normal to the surface may be discontinuous across the edges separating triangles: although at such points the curvature becomes, formally, infinite, their contribution to the integral vanishes. Specifically, if the normals to boundary points on two adjacent triangles form a nonzero angle α , and we model the sharp edge as a rounded edge of curvature radius $\rho \rightarrow 0$, the resulting path length is $\Delta s = \alpha \rho$, and the contribution to the integral behaves as $\Delta s / \rho^{2/3} = \alpha \rho^{1/3} \rightarrow 0$.
3. The direction (tangent vector) $\mathbf{t}_\tau$ of the ray (geodesic) on the face, again defined at the entry point $\mathbf{r}_\tau$. In addition, we define coefficients of the approximate quadratic parameterization $\mathbf{u}(s)$ of the geodesic trajectory in the parameter space $\mathbf{u}$. These coefficients are used in interpolating the quantities associated with the ray within the triangle.

4. The incident field projections

$$E_s^i = \mathbf{b}_0 \cdot \mathbf{E}^i(\mathbf{r}_0) \quad \text{and} \quad E_h^i = \mathbf{n}_0 \cdot \mathbf{E}^i(\mathbf{r}_0) \quad (7.28)$$

on the bi-normal and normal vectors at the shadow boundary point $\mathbf{r}_0$. These parameters are kept constant during the creeping ray evolution.

5. The width w_τ of the ray strip, which is also being updated during the ray evolution procedure by solving the geodesic-deviation equation, as mentioned in Sec. 7.3. In the scheme in which Since we are considering an infinitesimally narrow ray strip, the scale of $w(s)$ is arbitrary and immaterial, since only ratios of widths are relevant; we set $w(s_0) = 1$. In addition to the width value $w_\tau = w(s_\tau)$ we store coefficients of the approximate quadratic parameterization of the widths w , as a function of the geodesic length s on the considered triangle.

The parameters listed above are sufficient to determine the quantities $\mathbf{p}_{\tau m}$ and $\mathbf{q}_{\tau m}$ in the parameterization (7.27) (in the following we drop the index m). The simplest approximation is to identify the reference point $\mathbf{r}_{0\tau}$ with the ray entry point $\mathbf{r}_\tau$. In this approximation we take

$$\mathbf{q}_\tau = k \mathbf{t}_\tau \quad (7.29)$$

and

$$\mathbf{p}_\tau = [\mathbf{b}(\mathbf{r}_\tau) E_s^i \mathcal{T}_s(\mathbf{r}_\tau, \mathbf{r}_0) + \mathbf{t}(\mathbf{r}_\tau) E_h^i \mathcal{T}_h(\mathbf{r}_\tau, \mathbf{r}_0)] Q(s_\tau, s_0), \quad (7.30)$$

where the values of the functions $\mathcal{T}_s$, $\mathcal{T}_h$, and Q are expressible in terms of s_τ , ξ_τ , and w_τ .

A better approximation would involve locating $\mathbf{r}_{0\tau}$ at the triangle centroid, and interpolating the computed quantities to that point; our parameterization of the surface and the geodesics allows in this case the second-order interpolation. This interpolation can be then also used in conjunction with the second-order expansion of the phase factor in the representation (7.27). In this case it is more convenient to express the phase directly in the parameter space of the surface, i.e., in terms of the variables $(u^1, u^2) = \mathbf{u}$. For this purpose we have constructed a quadratic expansion of the phase factor and a similar (but linear) expansion of the vector $\mathbf{p}_\tau$ (these elements of the algorithm were developed, in a different context, under a separate SBIR contract).

The contribution of the given ray strip to the current on the triangle T_τ can be then parameterized as

$$\mathbf{J}_\tau(\mathbf{r}(\mathbf{u})) = [\mathbf{a}_\tau + \mathbf{b}_{\tau i} (u^i - u_0^i)] \exp \left\{ i \left[A_\tau + B_{\tau i} (u^i - u_0^i) + C_{\tau ij} (u^i - u_0^i) (u^j - u_0^j) \right] \right\} \quad (7.31)$$

where we assume the convention of summing over repeated indices $i, j = 1, 2$, and where $\mathbf{u}_0 = (1/3, 1/3)$ is the center of the unit triangle in the parameter space. The coefficients $\mathbf{a}$, $\mathbf{b}$, A , B , and C in the expansion are expressible in terms of the quadratic parameterization of the triangle geometry, and the ray parameters listed in the points 1–5 above (including, in this case, the quadratic parameterizations of the ray trajectory and of the geodesic deviation).

The main purpose of the current parameterizations (7.27) or (7.31) is to allow a fast computation of the scattered field, as a sum of Fourier transforms of these expressions. For instance, in the practical implementation of the Fourier transform algorithm for the representation (7.31) for a curved triangle, we subdivide the the unit triangle in the $\mathbf{u}$ space into triangles sufficiently small to be able to retain only the linear term in the exponent, and expand the remainder of the phase function to quadratic terms (higher order expansion is, of course, possible, but becomes rather complex). An additional element of the computation is an expansion of the Jacobian of the transformation $\mathbf{u} \longrightarrow \mathbf{r}(\mathbf{u})$ (i.g., the determinant of the metric tensor), required in integrating in the parameter space $\mathbf{u}$.

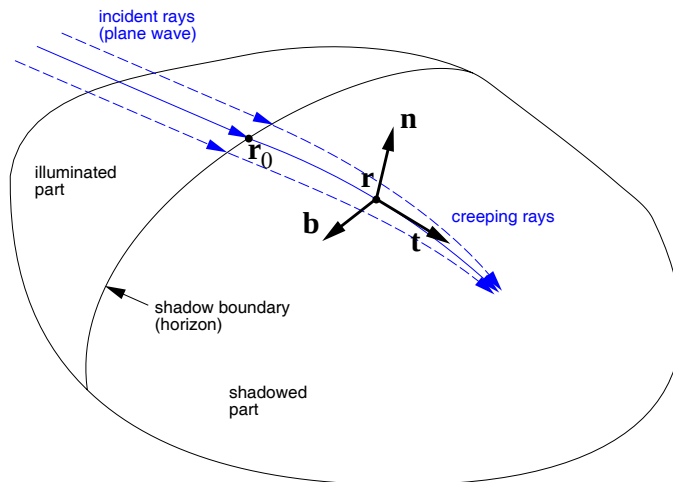


Figure 28: A creeping ray strip configuration used in surface current evaluation.

7.8 Implementation of algorithms for current computation

We now briefly discuss our implementation of the computational procedure described above. Although we have not parallelized the code, we mention in the following some parallelization aspects of the algorithm elements.

7.8.1 Computation of ray parameters

The parameters of the rays (points 1–5 in Sec. 7.7) are computed while performing ray tracing for rays emitted from the shadow boundary (that procedure is described in Sec. 7.4). For each triangle we store parameters of rays traversing that triangle. This algorithm is parallelizable in the sense that evolution of each ray strip (in a given “generation”) is independent. On the other hand, with the present algorithm for selecting distinct representative rays for each triangle requires either a shared or replicated storage for ray data or communication (if triangles are distributed across processors). An alternative procedure (mentioned in point 7 of Sec. 7.4) would be to collect data for all rays traversing all triangles, and select the optimal subset at the end of the computation.

7.8.2 Parameterization of the current

Parameterizations of the currents (such as (7.27) or (7.31)) are computed in a loop through triangles, after the reduced set of ray data becomes available. This procedure can be easily parallelized by distributing triangles across threads or processors.

7.8.3 Computation of the scattered field

The contribution to the scattered field is also computed for each triangle independently, and can be carried out in parallel in a trivial way.

7.8.4 Computation for multiple frequencies

The ray tracing procedure as described above is essentially independent of the frequency – the only dependence is through a ray termination criterion based on the Fock parameter value. Therefore, if the rays are traced sufficiently far to cover the lowest considered frequency, the constructed rays and current parameterization can be reused for multiple frequencies, provided the wave number and the Fock parameter are properly rescaled. This procedure can also be effectively applied to obtain scattered fields in time domain.

7.9 Applications of creeping-ray algorithms

As a means of validation of our approach to the creeping-wave formulation and its numerical implementation, we analyze the problem of a high-frequency wave scattering on three body-of-revolution (BOR) geometries:

1. a sphere,
2. a spheroid, and
3. a “cone-double-sphere” (a cone-sphere with a rounded cone tip).

In Fig. 29 we show the second and the third object, with the current distribution computed using the rigorous (MoM) solution.

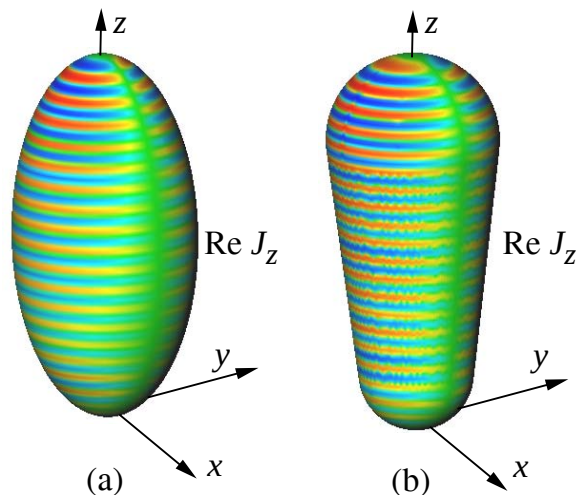


Figure 29: Currents on the spheroid (a) and cone-double-sphere (b) objects, computed with the MoM solver for the plane wave of vertical polarization incident along the negative z -axis ($\theta_i = \phi_i = 0^\circ$). The radii of the objects are approximately $R \simeq 4.54 \lambda$.

The reason for this particular choice of geometries is that the first and the third object, when illuminated along the axial direction, are characterized by the same PO return, but by significantly different behavior of creeping waves in the shadow region, especially on the conical segment of the third object, with the vanishing surface curvature along the creeping waves propagation direction. The spheroid geometry (although its PO return is different), provides an intermediate case with a gradually varying longitudinal curvature.

Another motivation was related to the recently developed Shadow Boundary Incremental Length Diffraction Coefficients (SBILDCs) technique [11]. That approach, while efficient and accurate in many three-dimensional problems, is sensitive only to the geometry properties *at* the shadow boundary (SB): the shape of the SB curve, the surface normal at the SB, and the curvature along the geodesics at the SB. Therefore, in our examples, the SBILDC approach would predict the same bistatic cross-section in the sphere and cone-double-sphere cases, while the rigorously computed cross-sections are quite different. We show in Fig. 30 the comparison of the vertical- (E -plane) bistatic cross-sections; generally, the horizontal-polarization cross-sections are less structured, and the differences are smaller. It is seen, in

particular, that the frequency of oscillations in the cross-section appears to depend on the longitudinal size of the object, and hence on the current distribution in the shadow region.

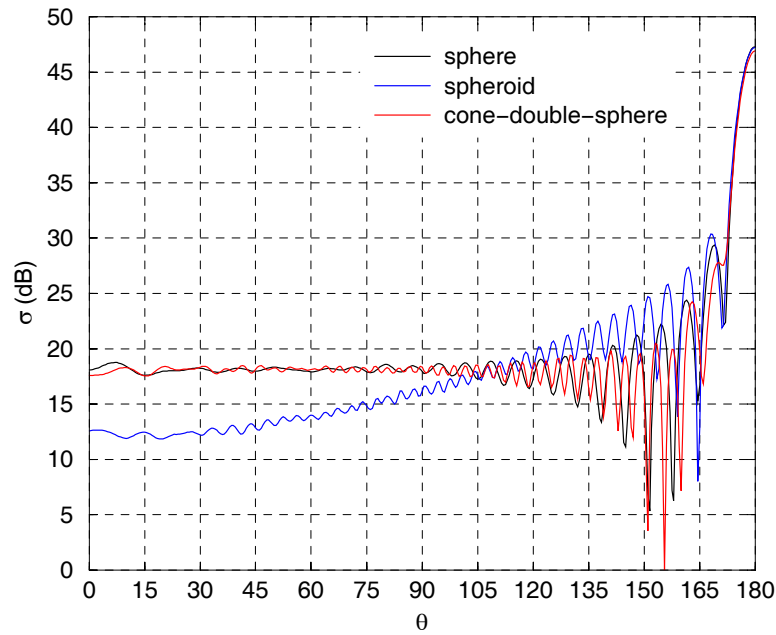


Figure 30: The vertical- (E -plane-) polarization bistatic cross-section for the axially incident plane wave on the perfectly conducting sphere, the spheroid, and the cone-double-sphere, computed with the rigorous (MoM) solution.

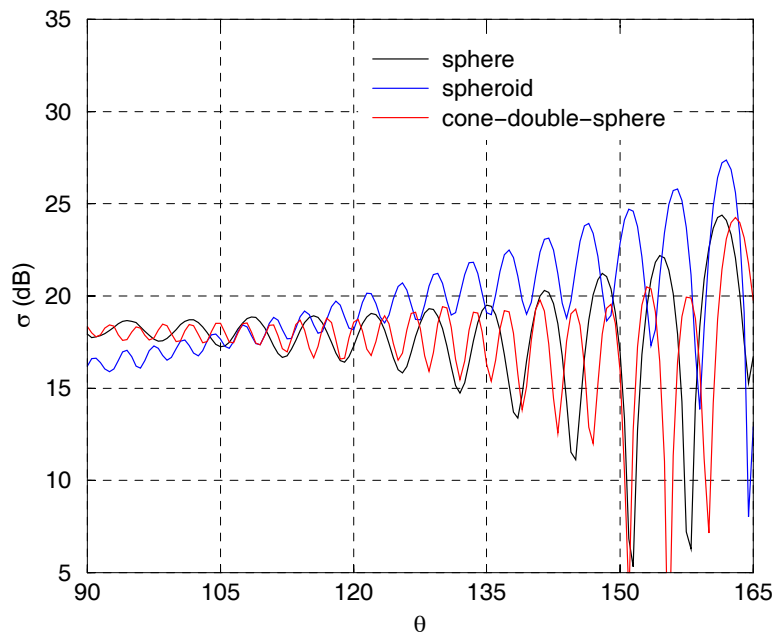


Figure 31: The cross-section of Fig. 30 shown in a limited angular range, exhibiting different oscillatory patterns.

In what follows we present results obtained with:

1. Rigorous (MoM) computations.
2. Semi-analytic computations of currents due to creeping rays and modified PO currents (as described in Sec. 7.6). In the case of BOR type geometries, it is possible to reduce the computation of the scattered field to a single numerical integral along the generating curve of the BOR; the remaining calculations can be carried out analytically. We parameterize the surface of the BOR, with the axis along the z -axis, as $(r, z) = (r(s), z(s))$, where r is the distance from the z -axis, and s is the distance measured along the surface, starting with $s = 0$ on the symmetry axis in the illuminated area.
3. Our numerical approach of constructing creeping waves and parameterizing surface current described in Secs. 27.5 and 7.8.

7.9.1 Semi-analytic construction of creeping rays for a body of revolution

1. Scattering on a sphere

As the first example we consider bistatic scattering on a perfectly conducting sphere of radius $R = 2.974$ in, at frequency $f = 18$ GHz, corresponding to $\lambda \simeq 0.655$ in and $R \simeq 4.54 \lambda$. The size of the object is chosen to match a similar problem in the published set of benchmarks [12].

In Fig. 32 we compare results of the rigorous (Mie solution) computation, the Physical Optics (PO), and the PO supplemented with the creeping wave (Fock) currents, according to the procedure described above, for the vertical polarization. Again, in the horizontal-polarization case the differences are smaller. We took the incidence angles $\theta_i = \phi_i = 0^\circ$, and computed the bistatic scattering cross-section as the function of θ in the $\phi_i = 0^\circ$ plane, for the horizontal (H -plane) polarization (the incident electric field along the y -axis), and the vertical (E -plane) polarization (the incident electric field along the x -axis).

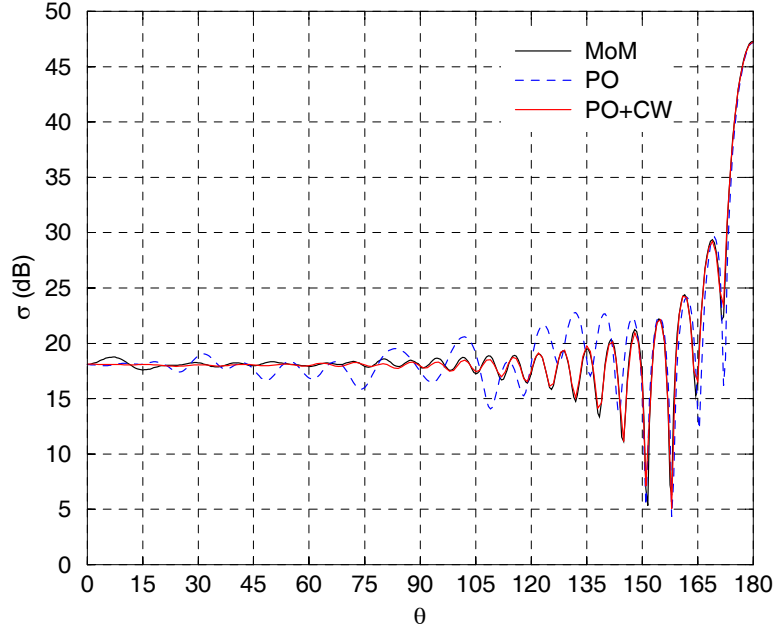


Figure 32: The vertical-polarization bistatic cross-section on the perfectly conducting sphere of $R \simeq 4.54 \lambda$, at the incidence angles $\theta_i = \phi_i = 0^\circ$: the rigorous result (Mie solution), Physical Optics (PO), and PO with creeping waves (PO+CW).

A known characteristic feature of the horizontal- and vertical-polarization problems for a sphere and similar bodies of revolution is a relatively smooth cross-section in the former and a more structured cross-section in the latter case. In both instances, the PO approximation yields cross-sections more oscillatory than for the rigorous result. However, after including the creeping-wave (CW) contributions, we find a good agreement with the MoM (or Mie) solution, except for some discrepancies (of about 1 dB) in near-forward scattering for horizontal polarization, and in near-backward scattering for vertical polarization.

These characteristics of the exact and approximate solutions can be interpreted in terms of the behavior of the radiating currents on the object surface, shown in Fig. 33. We plot there the “soft” and “hard” current components J_b and J_t , i.e., the components in the directions

of the binormal vector $\mathbf{b}$ and the tangential vector $\mathbf{t}$ of the geodesic (Eq.(7.13)).

The scattering amplitude is, essentially, related to the Hankel transforms of the currents $J_b(s)$ and $J_t(s)$. More specifically, it is expressible in terms of the integrals

$$\int_0^{s_{\max}} ds r(s) (\cdots) J_\nu(k r(s) \sin \theta) e^{-i k z(s) \cos \theta} J_t(s) , \quad (7.32)$$

where θ is the scattering angle, the coordinates $r = \sqrt{x^2 + y^2}$ and z parameterize the surface of the scatterer, the ellipses denote smoothly varying factors, and the Bessel functions J_ν appear with orders $\nu = 1, 2$. The geometry of the scattering process dictates that the soft currents dominate in the horizontal-polarization cross-section, and hard currents in the vertical-polarization cross-section.

The spike in the current amplitude at the south pole is due to the ray convergence factor $[w(s_0)/w(s)]^{1/2} \sim 1/r(s)^{1/2}$, and the current to the right of that point represents contributions of rays which have crossed the focal point. (We note, however, that the spike is suppressed in the integral of Eq.(7.32) by the Jacobian factor $r(s)$.)

In fact, the current distribution of Fig. 33 is in close agreement with the distributions computed from the Mie solution for the sphere; the deviations are practically limited only to the vicinity of the south pole, where the Mie solution currents are smaller and more smooth than predicted by the Fock formula.

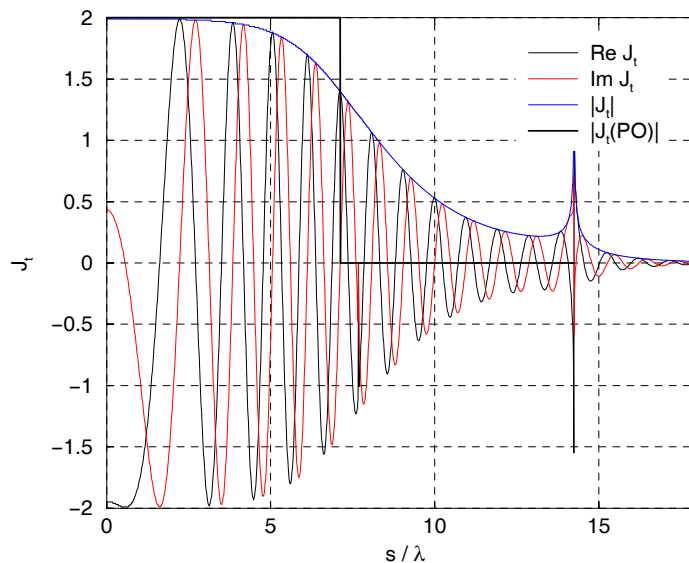


Figure 33: The hard current component J_t on the sphere, plotted as the function of s/λ . The absolute value of the PO current is also shown.

A comparison of the PO approximation currents and those including creeping-wave contributions indicates that the stronger oscillations in the PO cross-section is due to the less smooth termination of the PO currents at the shadow boundary (especially for the hard current J_t). The total currents (including the CW contributions) are, however, smooth across the shadow boundary, and the wider distribution of the hard current J_t (its deeper penetration into the shadow region) results in the more apparent oscillatory structure of the vertical-polarization cross-section in the medium angular range ($75^\circ \lesssim \theta \lesssim 150^\circ$).

This and similar examples show that the hard creeping wave contributions in the shadow region dominate the soft terms; on the other hand, modifications of the PO contributions are relatively more important in the soft currents.

2. Scattering on a spheroid

As the second example, we consider bistatic scattering on a perfectly conducting prolate spheroid, with the symmetry axis along the z -axes. The semi-minor axis of the spheroid is the same as that of the previously considered sphere ($R = 2.974$ in), and the semi-major axis is

twice as long ($R' = 5.948$ in). We assume axial plane wave incidence at the frequency $f = 18$ GHz, corresponding to the object length of about 10.5λ .

Fig. 34 compares bistatic cross-sections for the rigorous MoM computation, the PO, and the PO with the creeping wave currents. (The MoM computation, carried out with a code for arbitrary three-dimensional geometries, required about 100,000 unknowns.) The general features of the cross-sections are similar to those for the sphere, but the oscillations (better visible in the vertical-polarization case) are more rapid, due to the larger longitudinal size of the object.

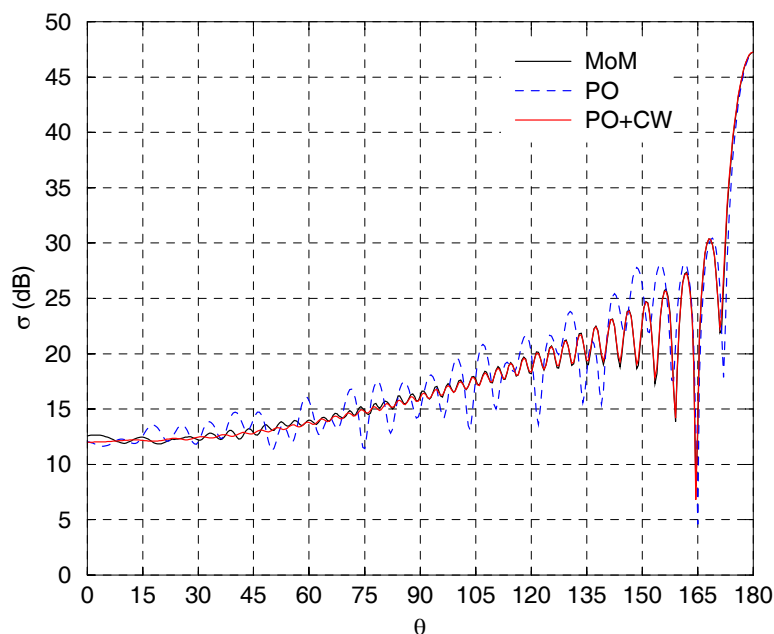


Figure 34: The vertical-polarization bistatic cross-section on the perfectly conducting prolate spheroid at 18 GHz, at the incidence angles $\theta_i = \phi_i = 0^\circ$: the rigorous result (MoM), PO, and PO with creeping waves (PO+CW).

The corresponding current distributions, shown in Fig. 35 for the hard component, indicate that, due to the smaller surface curvature along the geodesics, the creeping waves penetrate into the shadow region deeper than in the case of the sphere, especially in the hard current

contribution. It is these currents that enable us to correctly reproduce oscillations in the cross-section such as shown in Fig. 34.

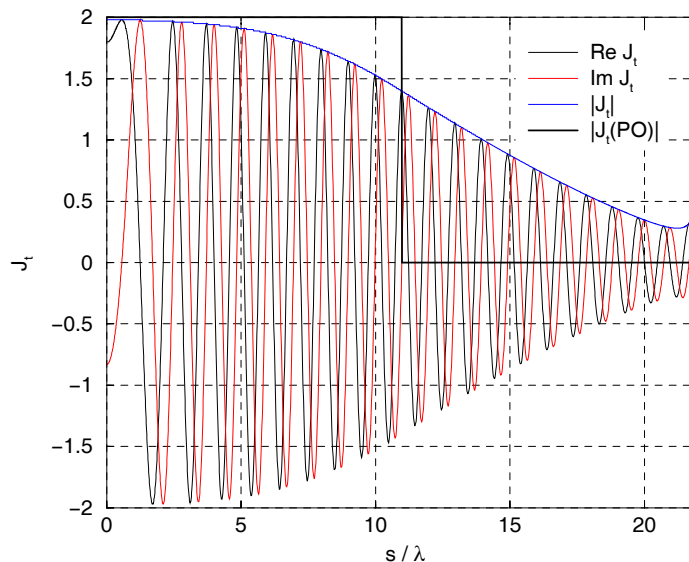


Figure 35: The hard current component J_t on the spheroid, plotted as the function of s/λ , shown together with the absolute value of the PO current.

3. Scattering on a cone-double-sphere

As the third example, we analyze bistatic scattering on a perfectly conducting “cone-double-sphere” object shown in Fig. 29 – a modification of the cone-sphere target of Ref. [12], in which the larger sphere has the radius $R = 2.974$ in, and the conical tip has been rounded with another sphere of radius $R' = 1.974$ in, such that both sphere surfaces are tangent to the cone surface. As in the original object, the cone opening half-angle is 7° , and, as before, the incident wave frequency is $f = 18$ GHz. The length of the object is then $L \simeq 20 \lambda$.

The main difference between the previous examples and the present object is that now the curvatures vary strongly along the object length. In particular, for axial incidence, the creeping rays propagate in planes passing through the object axis of symmetry; and in this case the curvature of the rays along the geodesic, $1/\rho_g$, vanishes on the conical surface. Therefore, according to the original Fock theory, the radiat-

ing currents are (practically) zero on the conical part of the object, and the radiation comes only from a rather narrow strip around the shadow boundary, and from the rounded back end of the object. The expected effect of this current distribution is an oscillatory behavior of the bistatic cross-section, with the oscillation frequency characteristic of the object length ($L \simeq 20 \lambda$); these oscillations should be about twice as rapid as those for the illuminated sphere (diameter $2R \simeq 9.1 \lambda$).

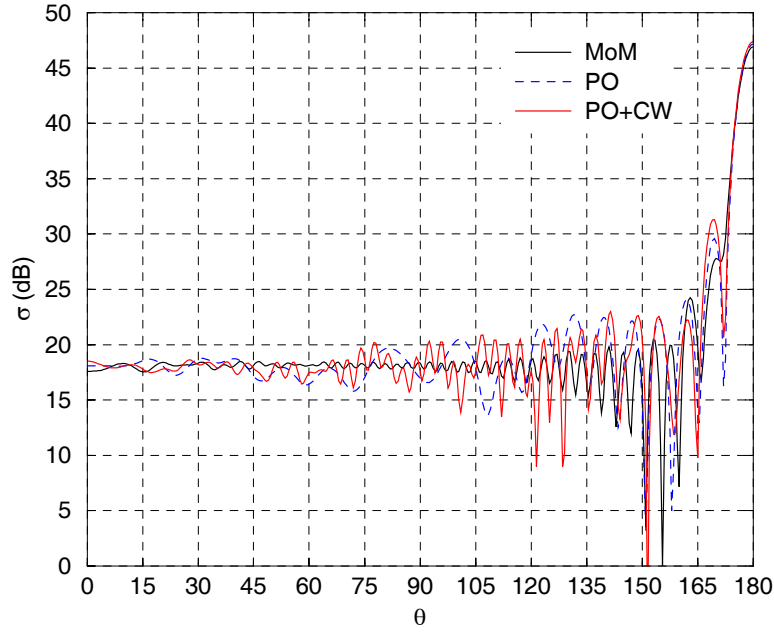


Figure 36: The vertical-polarization bistatic cross-section on the perfectly conducting cone-double-sphere at 18 GHz, at the incidence angles $\theta_i = \phi_i = 0^\circ$: the rigorous result (MoM), PO, and PO with creeping waves (PO+CW).

In Fig. 36 we compare results for the rigorous MoM computations, the PO, and the PO with the creeping wave currents.

In this problem the agreement between the MoM solution and the PO with creeping waves is much worse than for the previous cases:

- (a) In the horizontal-polarization case (not shown here), although the PO+CW cross-section reproduces on the average the level of the rigorous solution, it exhibits large-amplitude oscillations for an-

gles near the forward scattering direction, and smaller amplitude oscillations for smaller angles. Both of these types of oscillations are absent in the rigorously computed cross-section.

- (b) In the vertical-polarization case, the PO+CW cross-section for near forward-scattering angles follows rather closely the PO approximation. For smaller angles, although the rapid oscillations of the PO+CW result approximately match the oscillatory pattern of the exact cross-section, the PO+CW cross-section appears to also include a less rapidly oscillating component, similar in frequency and amplitude the PO contribution.

The features of the PO+CW cross-section can be interpreted in terms of the currents distributions, plotted in Fig. 33 for the hard currents. The creeping rays contributions in the shadow region extend from $s \simeq 7.12\lambda$ to $s \simeq 7.67\lambda$ (from the shadow boundary to the sphere-cone transition point), and from $s \simeq 20.09\lambda$ to $s \simeq 24.44\lambda$ (the beginning of the smaller sphere surface to the south pole). There is no radiating current on the conical surface, since in this area the curvature in the geodesic direction is zero (the factor $[\rho_g(s_0)/\rho_g(s)]^{1/6}$ in Eq.(7.13)).

The rapid oscillations in the PO+CW cross-sections of Fig. 36 are due to the interference of fields radiated by the currents on the two spherical surfaces. This feature appears to be correct, since the MoM cross-section also exhibits such oscillations (more visible in the vertical polarization). On the other hand, the Fock current distribution of Fig. 37 is unable to compensate for the *incorrect* oscillations of the cross-section present in the PO approximation. The reason may be that, unlike in the case of a sphere (Fig. 33), the contribution of the shadow region creeping waves on the larger sphere is small (it comes only from a thin ring between the shadow boundary and the boundary between the sphere and the cone). At the same time, the contribution of the creeping waves on the smaller sphere is not able to compensate the smooth oscillations in the PO approximation without introducing at the same time large-amplitude rapid oscillations due to the interference.

We find thus that the original Fock formulation does not provide sufficient accuracy in situations where the surface curvature is discontinuous, and, in particular, when the curvature along the geodesic is much smaller (as in our case, where it vanishes) than the curvature in the transverse direction. There are indications that the transverse curvature does play a role in such situations [13].

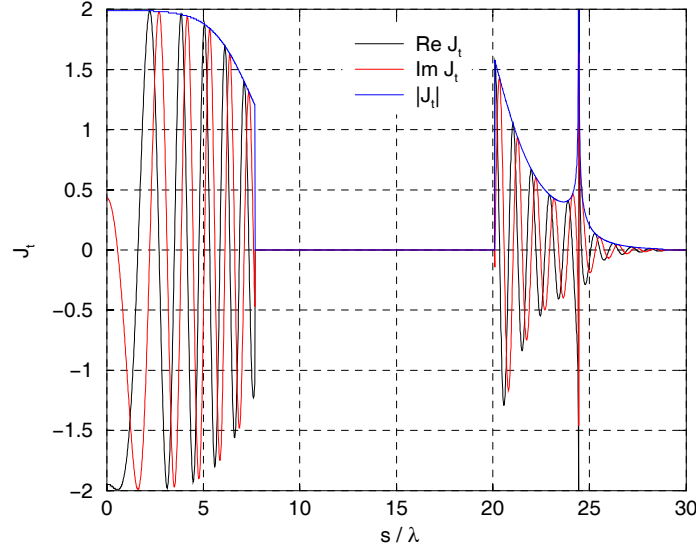


Figure 37: The hard current component J_t on the cone-double-sphere, plotted as the function of s/λ .

7.9.2 Numerical ray tracing and current computation

1. Scattering on a sphere

As before, we consider bistatic scattering on a perfectly conducting sphere of radius $R = 2.974$ in, at frequency $f = 18$ GHz, corresponding to $\lambda \simeq 0.655$ in and $R \simeq 4.54 \lambda$.

We consider, as previously, the axial incidence, and assume a vertically polarized incident plane wave. For the horizontal polarization the results are similar, although the creeping-wave effects are smaller.

In a typical computation of this type, according to the algorithms of Sec. 7.5, we generate about 300 geodesics, and the entire computation, including evaluation of the cross-section, takes less than 2 minutes.

In Fig. 38 we compare results of the rigorous MoM computation, the Physical Optics (PO), and the PO supplemented with the creeping wave (Fock) currents, according to the numerical procedure described in Secs. 7.5 and 7.8.

The agreement is comparable to that obtained with the semi-analytic formulation, and the oscillatory structure of the cross-section (largely

due to the behavior of currents in the shadow region) correctly in most of the angular range (for $\theta > 60^\circ$). We find that the oscillations in the cross-section are highly sensitive to the phase of the creeping wave, in particular to the phase variation of the Fock function $g(\xi)$ (Eqs. 7.19 and 7.22). In this and similar problem we also find that the hard creeping wave contributions dominate the soft terms.

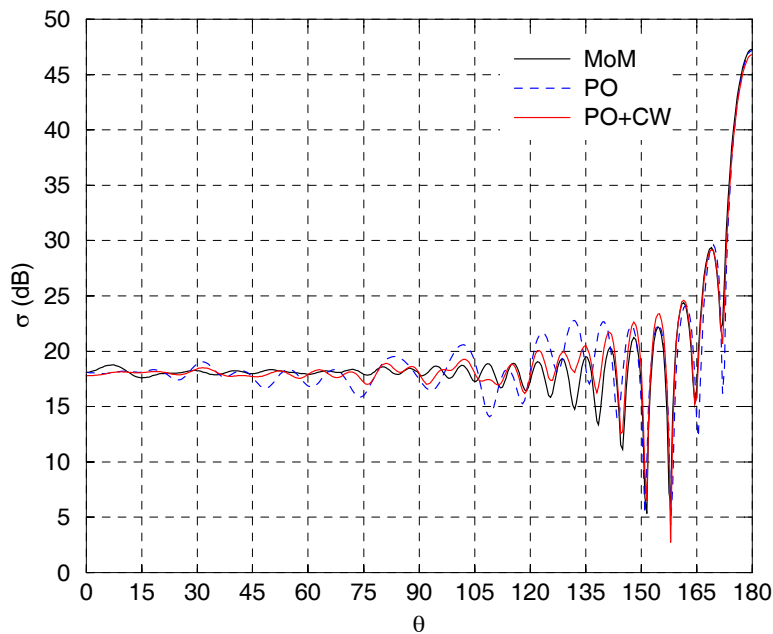


Figure 38: The vertical-polarization bistatic cross-section on the perfectly conducting sphere at 18 GHz, at the incidence angles $\theta_i = \phi_i = 0^\circ$: the rigorous result (MoM), (PO), and **numerically** evaluated asymptotic solution with creeping waves (PO+CW).

2. Scattering on a spheroid

The results of the numerical computation of the creeping waves' contributions to the bistatic cross-section in this case are shown in Fig. 39 for the (more difficult) case of the vertical polarization. Again, in most of the angular range, the agreement between the MoM and the numerical creeping waves' result is nearly as good as for the semi-analytic formulation; the important feature of the numerical computation is

that it correctly reproduces the oscillations in the cross-section, due mostly to the behavior of the solution in the deep-shadow regions.

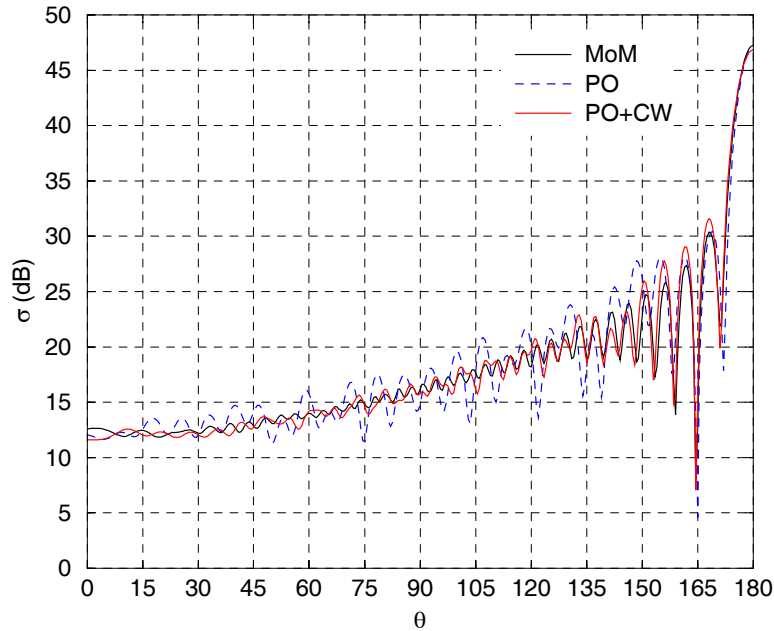


Figure 39: The vertical-polarization bistatic cross-section on the perfectly conducting spheroid at 18 GHz, at the axial incidence: the rigorous result (MoM), PO, and *numerically* evaluated asymptotic solution with creeping waves (PO+CW).

3. **Scattering on a cone-double-sphere** In the case of the cone-double-sphere, the numerical results are in good agreement with the semi-analytic calculation (Fig. 36), although neither of them agrees with the rigorous MoM result, due to the deficiencies of the asymptotic theory in the case of a discontinuous surface curvature. (as discussed in Sec. 7.9.1).

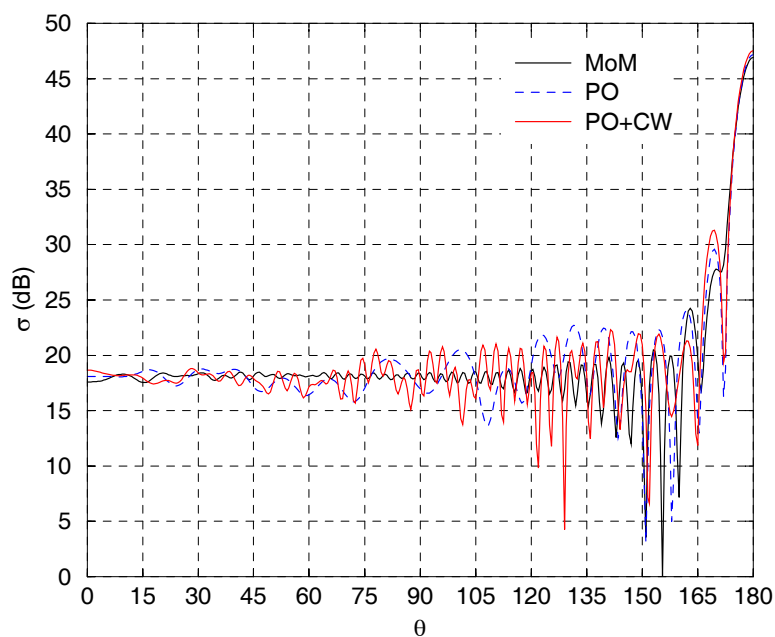


Figure 40: The vertical-polarization bistatic cross-section on the perfectly conducting cone-double-sphere at 18 GHz, at the axial incidence: the rigorous result (MoM), PO, and the *numerical* computation of the asymptotic solution with creeping waves (PO+CW).

References

- [1] R.J. Burkholder and P.H. Pathak, "Geodesic Ray tracking on triangulated surfaces", The Ohio State University report No.745773-2, June 2004.
- [2] P. Janpugdee, P.H. Pathak and K. Tap, "A novel DFT-based UTD ray analysis for the collective radiation from large finite cylindrical conformal arrays and apertures", The Ohio State University report No.745773-1, June 2004.
- [3] G.M. Nielson, "A transfinite, visually continuous, triangular interpolant", in: G.E. Farin (ed.), *Geometric modeling: algorithms and new trends*, SIAM, Philadelphia, pp. 235-246, 1987.
- [4] H. Hagen and H. Pottman, "Curvature continuous triangular interpolants", in: T. Lyche and L.L. Schumaker (eds.), *Mathematical methods in Computer Aided Geometric Design*, Academic Press, San Diego, pp. 373-384, 1989.
- [5] G.A. Deschamps, "Ray techniques in electromagnetics", *Proceedings of the IEEE*, Vol. 60(9), pp. 1022-1035, 1972.
- [6] D. Bouche and f. Molinet, *Methodes Asymptotiques en Electromagnetisme*, Springer-Verlag, 1994.
- [7] V.A. Fock, "The field of a plane wave near the surface of a conducting body", *J. Phys.*, Vol. 10, pp. 399-409, 1946.
- [8] I. Andronov and D. Bouche, "Friedlander-Keller solution for 3-d Maxwell case", in *Progress in Electromagnetics Research, PIER 15*, pp. 87-108, 1997.
- [9] T.B. Hansen and R.A. Shore, "Incremental length diffraction coefficients for the shadow boundary of a convex cylinder", *IEEE Trans. Antenn. Propagat.*, Vol. AP-46, pp. 1458-1466, 1998.
- [10] G.L. James, *Geometrical Theory of Diffraction for Electromagnetic Waves*, Peter Peregrinus Ltd., London, U.K., 1986.
- [11] R.A. Shore and A.D. Yaghjian, "Shadow boundary incremental length diffraction coefficients applied to scattering from 3-d bodies", *IEEE Trans. Antenn. Propagat.*, Vol. AP-49, pp. 200-210, 2001.

- [12] A.C. Woo, H.T.G. Wang, M.J. Schuh, and M.L. Sanders, "Benchmark radar targets for the validation of computational electromagnetic programs", *IEEE Antennas and Propagation Magazine*, Vol. 35, pp. 84-89, 1993.
- [13] I. Andronov, "Creeping waves on elongated bodies", in *State of the art in high frequency asymptotics and hybrid methods*, to appear.
- [14] V.H. Weston, "The effect of a discontinuity in curvature in high-frequency scattering", *IRE Trans. Antennas Propagat.*, Vol. AP-10, pp. 775-780, 1962.
- [15] V.H. Weston, "Extension of Fock theory for currents in the penumbra region", *Radio Sci.*, Vol. 69D, p. 1257-1270, 1965.
- [16] T.B.A. Senior, "The diffraction matrix for a discontinuity in curvature", *IEEE Trans. Antenn. Propagat.*, Vol. AP-20, pp. 326-333, 1972.
- [17] L. Kaminetsky and J.B. Keller, "Diffraction coefficients for higher order edges and vertices", *SIAM J. Appl. Math.*, Vol. 22, pp. 109-134, 1972.
- [18] G.L. James, G. Tong, and D.A. Ross, "Uniform diffraction solution for a discontinuity in curvature", *Electronics Letters*, Vol. 11, pp. 557-559, 1975.
- [19] A. Michaeli, "Diffraction by a discontinuity in curvature including the effect of the creeping wave", *IEEE Trans. Antenn. Propagat.*, Vol. AP-38, pp. 929-931, 1990.
- [20] W. Franz and K. Klante, "Diffraction by surfaces of variable curvature", *IEEE Trans. Antenn. Propagat.*, Vol. AP-7, pp. S68-S70, 1959.
- [21] J.B. Keller and B.R. Levy, "Decay exponents and diffraction coefficients for surface waves of non-constant curvature", *IRE Trans. Antennas Propagat.*, Vol. AP-7, pp. S52-S61, 1959.
- [22] S. Hong, "Asymptotic theory of electromagnetic and acoustic diffraction by smooth convex surfaces of variable curvature", *J. Math. Phys.*, Vol. 8, pp. 1223-1332, 1967.
- [23] D.R. Voltmer, "Diffraction by doubly curved convex surfaces", Ph.D. Dissertation, Ohio State University, Columbus, Ohio, USA.

8 Yale University Subcontract to Monopole Research, Topic # AF01-009

Our research under this contract addressed five issues, as follows.

1. Construct a two-dimensional version of the algorithm, and use it to conduct preliminary numerical experiments
2. Conduct detailed asymptotic analysis in two dimensions, with the goal of obtaining tight estimates on the error of the method at (reasonably) high frequencies
3. Depending on the outcome of two-dimensional experiments, start implementation in three dimensions.
4. Extend the existing rudimentary "fast" solvers to Lippman-Schwinger equations in two dimensions.

While the work on the fast solvers continued more or less according to plan, our investigation into asymptotics produced a number of unexpected observations. Connections have been discovered to algorithms for the application to arbitrary vectors of certain matrices related to classical Sturm-Liouville operators and (not unexpectedly) to several types of classical special functions. Some of these appear to have independent technical value, and to deserve publication on their own merits; in addition, junior members of the group have participated significantly in this work, and should not be denied rapid conclusion and publication of these investigations. Thus, it seemed appropriate to tolerate a delay in some of the activities stipulated above, in order to bring these results to prompt publication. Specifically, the work on p. 2 above has been sidetracked, and the work on point 3. has been retarded. Below, some of our results are summarized in two reports; the first of these has been submitted for publication, and the second has been accepted at the Journal of Computational physics.

9 Approximate Formulae For Certain Prolate Spheroidal Wave Functions Valid for Large Values of Both Order and Band-Limit

9.1 Introduction

Originally, Prolate Spheroidal Wave Functions were discovered as the eigenvectors of the differential operator D_c defined by the formula

$$D_c(\psi)(x) = (1 - x^2) \psi''(x) - 2x \psi'(x) - c^2 x^2 \psi(x) \quad (9.1)$$

subject to the condition that ψ is continuous on the interval $[-1, 1]$. For each positive real c , there exists a countable set of real numbers $\chi_0 < \chi_1 < \chi_3 < \dots$ for which the equation

$$(1 - x^2) \psi''(x) - 2x \psi'(x) + (\chi_j - c^2 x^2) \psi(x) = 0 \quad (9.2)$$

has a continuous solution on $[-1, 1]$. Such coefficients are known as prolate eigenvalues, and the corresponding solutions of (9.2) are referred to as Prolate Spheroidal wave Functions (PSWFs). About 45 years ago, it was observed (see [12, 7, 10]) that PSWFs are also eigenvectors of the integral operator

$$F_c(\psi)(x) = \int_{-1}^1 e^{icxt} \psi(t) dt; \quad (9.3)$$

it is in this latter capacity that the PSWFs are of interest to the authors of the present paper, due to the obvious connections between (9.3) and Fourier Transforms, Toeplitz matrices, antennas, Heisenberg principle, etc.

Despite their remarkable role as eigenvectors of (9.3), PSWFs have not been investigated in as much detail as several other classes of special functions; the reasons seem to be related to the fact that the classical scheme for their numerical evaluation (the so-called Bouwkamp method) tends to fail for large values of c . In [15], we observe that a simple modification of the Bouwkamp scheme converts it into a reliable numerical tool for virtually all values of c that are likely to be encountered in practice; we also summarize a number of analytical properties of PSWFs, vaguely reminiscent of the properties of Bessel Functions. In [16], we investigate the asymptotic behavior of ψ_m^c for large c and fixed m .

For any given c , PSWFs constitute an orthonormal basis on the interval $[-1, 1]$; numerical evidence is accumulating that in certain situations, they

are preferable as a numerical tool to classical polynomial bases (such as Legendre and Chebychev polynomials). On the other hand, it has been observed that when PSWFs have to be dealt with analytically (i.e. when the time comes to put proofs behind facts observed numerically), the analysis tends to be much harder than that encountered while dealing with most classical special functions (orthogonal polynomials, Bessel functions, spherical harmonics, etc.). In the experience of the authors of the present paper, this relative difficulty has more to do with the paucity of published results about the PSWFs, rather than with the inherent difficulty of the subject. In this paper, we investigate PSWFs and the corresponding eigenvalues χ_m^c , λ_m^c in the regime when $m \geq c$; while most of the properties we derive are more or less obvious for $m \gg c$, the behavior of PSWFs is considerably more subtle when $m > c$, but m/c is not very large. The paper is meant to be a compendium of properties of PSWFs and the corresponding eigenvalues χ_m^c , λ_m^c that the authors found to be useful in their attempts to utilize PSWFs as a numerical tool. While most of the material presented here appears to be new, no serious effort has been made to separate original results from those published previously.

This paper contains two types of results. The first kind are expansions of various quantities (PSWFs, corresponding eigenvalues, etc.) into powers of c/m , valid when the ratio c/m is small (or, in some cases, not very large). Most of these expansions are of fairly high order (from 8 to 12), and have been obtained by the analysis of the three-term recursion connecting the coefficients of the Prolate expansion of a function with the coefficients of the Legendre expansion of the same function (see Theorem 3 and Observation 4 below). Once the formulae in Section 9.10 are obtained, each of them is easily verified by substituting it into (9.2), and using the identities (9.7), (9.4), (9.13). This approach is very similar to that used in [16] to obtain high-frequency asymptotics for PSWFs and their corresponding eigenvalues.

The second type of results presented below are asymptotic expansions of PSWFs and corresponding eigenvalues for large m and c , presented in Section 9.13. These expansions are of low order (from $1/\sqrt{m \cdot c}$ to $1/m^2$), and obtained via straightforward WKB analysis of equation (9.2). While the derivation of the formulae in Section 9.13 is straightforward and uses classical techniques (see, for example, [2]), it is quite detailed, and will be published at a later date; the results are included here for completeness.

The paper is organized as follows. In Section 9.2, we summarize a number of well-known mathematical facts to be used in this paper. In Section 9.6 we introduce the analytical apparatus used to derive the asymptotics pre-

sented in the paper. Section 9.11 is devoted to asymptotic formulae for eigenvalues and eigenvectors of (9.1), and in Section 9.12, we construct such formulae for eigenvalues of (9.3). Section 9.13 contains a different type of asymptotic formulae (see the preceding paragraph), and the numerical behavior of some of the presented approximations is illustrated in Section 9.14. Finally, Section 9.15 contains generalizations and conclusions.

9.2 Mathematical Preliminaries

In this section, we introduce notation and summarize several well-known facts to be used in the rest of the paper.

9.3 Legendre Polynomials

In agreement with standard practice, we will be denoting by P_n the classical Legendre polynomials, defined by the three-term recursion

$$P_{n+1}(x) = \frac{2 \cdot n + 1}{n + 1} \cdot x \cdot P_n(x) - \frac{n}{n + 1} \cdot P_{n-1}(x), \quad (9.4)$$

with the initial conditions

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x; \end{aligned} \quad (9.5)$$

as is well-known,

$$P_k(1) = 1 \quad (9.6)$$

for all $k = 0, 1, 2, \dots$, and each of the polynomials P_k satisfies the differential equation

$$(1 - x^2) \cdot \frac{d^2 P_k(x)}{dx^2} + 2 \cdot x \cdot \frac{d P_k(x)}{dx} + k \cdot (k + 1) \cdot P_k(x) = 0. \quad (9.7)$$

The following two lemmas summarize several well-known facts about Legendre polynomials. All of these facts can be found, for example, in [5].

Lemma 1. *For any positive integer m ,*

$$\int_{-1}^1 x^m P_m(x) dx = \frac{\sqrt{\pi} m!}{2^m \Gamma(m + \frac{3}{2})}, \quad (9.8)$$

and for any complex z ,

$$P_m^{(m)}(z) = \frac{(2m)!}{2^m m!}. \quad (9.9)$$

Lemma 2. *Suppose that k, n are non-negative integers, $n \leq k$, and $k - n$ is even. Then*

$$\left| \int_{-1}^1 x^k P_n(x) dx \right| = \frac{\sqrt{\pi} k!}{2^k \left(\frac{k-n}{2}\right)! \Gamma(k + \frac{3}{2})}. \quad (9.10)$$

If $n > k$ or $k - n$ is odd,

$$\int_{-1}^1 x^k P_n(x) dx = 0. \quad (9.11)$$

The polynomials defined by the formulae (9.4), (9.6) are orthogonal on the interval $[-1, 1]$; however, they are not orthonormal, since for each $n \geq 0$,

$$\int_{-1}^1 (P_n(x))^2 dx = \frac{1}{n + 1/2}; \quad (9.12)$$

the normalized version of the Legendre polynomials will be denoted by $\overline{P}_n$, so that

$$\overline{P}_n(x) = \frac{P_n(x)}{\sqrt{n + 1/2}}. \quad (9.13)$$

Thus, $\overline{P}_n$ constitute an orthonormal basis in $L^2[-1, 1]$.

9.4 Elliptic Integrals

Incomplete elliptic integrals $F(x, a)$, $E(x, a)$ are defined by the formulae

$$F(x, a) = \int_0^x \frac{dt}{\sqrt{1 - a^2 \cdot \sin^2(t)}}, \quad (9.14)$$

$$E(x, a) = \int_0^x \sqrt{1 - a^2 \cdot \sin^2(t)} dt, \quad (9.15)$$

respectively. Complete elliptic integrals $F(a)$, $E(a)$ are defined by the formulae

$$F(a) = F(\pi/2, a) = \int_0^{\pi/2} \frac{dt}{\sqrt{1 - a^2 \cdot \sin^2(t)}}, \quad (9.16)$$

$$E(a) = E(\pi/2, a) = \int_0^{\pi/2} \sqrt{1 - a^2 \cdot \sin^2(t)} dt. \quad (9.17)$$

We will denote by $\tilde{E}$ the function inverse to E , so that

$$\tilde{E}(E(a)) = a, \quad (9.18)$$

and by $\tilde{G}$ the function defined by the formula

$$\tilde{G}(E(x, a), a) = x; \quad (9.19)$$

in other words, for a fixed a , the function $\tilde{G}$ is the inverse of $E(x, a)$ with respect to the argument x .

Finally, we will denote by

$$f = \overline{E}(x) \quad (9.20)$$

the solution of the equation

$$\int_0^{\pi/2} \sqrt{f - \sin^2(t)} dt = x \quad (9.21)$$

(viewed as an equation with respect to f), and observe the obvious connection between $\overline{E}$ and $\tilde{E}$.

9.5 Prolate Spheroidal Wave Functions

In this section, we summarize a number of analytical properties of the Prolate Spheroidal Wave Functions. Unless stated otherwise, all facts collected below can be found in [12, 7].

For any $c > 0$, we denote by F_c the operator $L^2[-1, 1] \rightarrow L^2[-1, 1]$ defined by the formula

$$F_c(\varphi)(x) = \int_{-1}^1 e^{icxt} \varphi(t) dt. \quad (9.22)$$

In other words, F_c is a finite Fourier integral operator depending on c , a parameter we frequently refer to as the *band-limit* of F_c .

Clearly, F_c is compact; we denote by $\lambda_0, \lambda_1, \dots, \lambda_n, \dots$ the eigenvalues of F_c in decreasing order such that $|\lambda_{n-1}| \geq |\lambda_n|$ for all natural n , and denote by ψ_n the corresponding eigenfunctions. In other words, for all non-negative integer n , λ_n and ψ_n satisfy the integral equation

$$\lambda_n \psi_n(x) = \int_{-1}^1 e^{icxt} \psi_n(t) dt \quad (9.23)$$

for all $x \in [-1, 1]$. In this paper, we adopt the convention that the functions are normalized such that $\|\psi_n\|_{[-1, 1]} = 1$ for all n . The following theorem is a combination of several lemmas from [12, 4, 6, 3].

Theorem 1. For any positive real c , the eigenfunctions $\psi_0, \psi_1, \dots$, of the operator F_c are purely real, are orthonormal, and are complete in $L^2[-1, 1]$. The even-numbered eigenfunctions are even, and the odd-numbered ones are odd. All eigenvalues of F_c are non-zero and simple; the even-numbered eigenvalues are purely real, and the odd-numbered ones are purely imaginary; in particular, $\lambda_n = i^n |\lambda_n|$. For each $i = 0, 1, 2, \dots$, the function ψ_i is an entire function of the two *complex* variables c, x .

We define the self-adjoint operator $Q_c : L^2[-1, 1] \rightarrow L^2[-1, 1]$ by the formula

$$Q_c(\varphi) = \frac{1}{\pi} \int_{-1}^1 \frac{\sin(c \cdot (x - t))}{x - t} \varphi(t) dt, \quad (9.24)$$

which relates to F_c in the form

$$Q_c = \frac{c}{2\pi} \cdot F_c^* \cdot F_c. \quad (9.25)$$

Consequently, Q_c has the same eigenfunctions as F_c , and the n -th (in descending order) eigenvalue μ_n of Q_c is given by the formula

$$\mu_n = \frac{c}{2\pi} \cdot |\lambda_n|^2. \quad (9.26)$$

The eigenvalues λ are analytic functions of c , and satisfy the differential equations

$$\frac{\partial \lambda_m}{\partial c} = \lambda_m \frac{2(\psi_m^c(1))^2 - 1}{2c}, \quad (9.27)$$

or, equivalently,

$$\frac{\partial \log(\lambda_m)}{\partial c} = \frac{1}{2c} (2(\psi_m^c(1))^2 - 1) \quad (9.28)$$

(see [3, 13]).

Obviously, the operator Q_c is closely related to the operator $P_c : L^2[-\infty, \infty] \rightarrow L^2[-\infty, \infty]$ defined by the formula

$$P_c(\varphi) = \frac{1}{\pi} \cdot \int_{-\infty}^{\infty} \frac{\sin(c \cdot (x - t))}{x - t} \cdot \varphi(t) dt, \quad (9.29)$$

which, as is well known, is the orthogonal projection operator onto the space of functions of band limit c on $(-\infty, \infty)$.

For large c , the spectrum of Q_c consists of three parts: about $2c/\pi$ eigenvalues that are very close to 1, followed by order $\log(c)$ eigenvalues which decay exponentially from 1 to nearly 0; the remaining eigenvalues are all very close to zero. More detailed discussions of the structure of the spectrum of Q_c can be found in [8], [15], and a number of other places.

By a remarkable coincidence, the eigenfunctions $\psi_0, \psi_1, \dots, \psi_n$ of the operator Q_c turn out to be the Prolate Spheroidal Wave Functions, well known from classical Mathematical Physics (see, for example, [9]). The following theorem formalizes this statement. It is proved in a considerably more general form in [10, 3].

Theorem 2. For any $c > 0$, there exists a strictly increasing sequence of positive real numbers $\chi_0, \chi_1, \dots$ such that for each integer $n \geq 0$, the differential equation

$$(1 - x^2) \psi''(x) - 2x \psi'(x) + (\chi_n - c^2 x^2) \psi(x) = 0 \quad (9.30)$$

has a solution that is continuous and bounded on the interval $[-1, 1]$. Moreover, for each integer $n \geq 0$, the function ψ_n (defined in Theorem 1) is the solution of (9.30).

9.6 Analytical Apparatus

In this section, we build a number of analytical tools for the construction of the formulae in this paper.

9.7 Legendre Coefficients and Three-Term Recursions

Prolate Spheroidal Wave Functions and Legendre polynomials are closely related, as can be observed from the similarities between their corresponding differential equations (9.30) and (9.7).

Since the scaled Legendre polynomials $\overline{P}_n$ constitute an orthonormal basis in $L^2[-1, 1]$, ψ_m^c can be expanded in the Legendre series

$$\psi_n^c(x) = \sum_{k=0}^{\infty} \beta_k^n \overline{P}_k(x), \quad (9.31)$$

with the coefficients β_k given by the formula

$$\beta_k = \int_{-1}^1 \psi_m^c(x) \overline{P}_k(x) dx; \quad (9.32)$$

since ψ_m^c is analytic on the interval $[-1, 1]$, the coefficients β_k^n decay exponentially once k is sufficiently large (a more detailed discussion of the rate of decay of $\{\beta_k^n\}$ can be found in [15]).

For small c , the connection between the functions ψ_m^c and Legendre polynomials $\overline{P}_m$ is well-known. *Inter alia*,

$$\psi_m^c(x) = \overline{P}_m(x) + O(c^2) \quad (9.33)$$

for all non-negative integer m and $x \in [-1, 1]$ (see, for example, [13]).

Substituting (9.31) into (9.30) and using (9.7), (9.4), (9.13), we obtain the well-known three-term recursion

$$\begin{aligned} & \frac{(k+2)(k+1)}{(2k+3)\sqrt{(2k+5)(2k+1)}} \cdot c^2 \cdot \beta_{k+2}^n + \\ & \left(k(k+1) + \frac{2k(k+1)-1}{(2k+3)(2k-1)} \cdot c^2 - \chi_n \right) \cdot \beta_k^n + \\ & \frac{k(k-1)}{(2k-1)\sqrt{(2k-3)(2k+1)}} \cdot c^2 \cdot \beta_{k-2}^n = 0. \end{aligned} \quad (9.34)$$

Introducing the notation

$$\beta^n = (\beta_0^n, \beta_1^n, \beta_2^n, \dots) \in l^2 \quad (9.35)$$

for all $n = 0, 1, 2, \dots$, we restate (9.34) in a slightly different form in the following theorem.

Theorem 3. Suppose that χ_n 's are eigenvalues of the differential operator (9.30), and that β^n 's are vectors defined in (9.35). Suppose further that a matrix A is given by the formulae

$$A_{k,k} = k(k+1) + \frac{2k(k+1)-1}{(2k+3)(2k-1)} \cdot c^2, \quad (9.36)$$

$$A_{k,k+2} = \frac{(k+2)(k+1)}{(2k+3)\sqrt{(2k+1)(2k+5)}} \cdot c^2, \quad (9.37)$$

$$A_{k+2,k} = \frac{(k+2)(k+1)}{(2k+3)\sqrt{(2k+1)(2k+5)}} \cdot c^2, \quad (9.38)$$

for all $k = 0, 1, 2, \dots$, the remainder of the entries being zero. Then, χ_n are the eigenvalues of A , with β^n corresponding eigenvectors.

In other words, in the basis consisting of the functions $\overline{P}_0, \overline{P}_1, \dots, \overline{P}_n, \dots$, the differential equation (9.30) has the form

$$(A - \chi_n \cdot I) \cdot \beta^n = 0. \quad (9.39)$$

Observation 4. The matrix A separates into two symmetric, tridiagonal matrices A_{even} , A_{odd} , with the former consisting of elements of A on the even-numbered rows and even-numbered columns, and the latter elements on odd-numbered rows and odd-numbered columns. While these two matrices are of infinite dimensions, and their entries do not decay much with increasing row or column numbers, the coordinates β_m^n of the eigenvectors β^n decay rapidly once m is sufficiently large.

9.8 Inverse Power Method as an Analytical Tool

In this section, we outline a scheme for the construction of asymptotic expansions in powers of c and $1/m$ for functions ψ_m^c and for certain quantities associated with them; the expansions constructed in this fashion are valid for large m (see Section 9.11 below). It must be pointed out that once expansions (9.64), (9.65), (9.67), (9.68) have been constructed, substituting them into (9.2) and using the identities (9.7), (9.4), (9.13) readily (though somewhat messily) proves their validity. Thus, this section should be viewed as a somewhat heuristic description of the procedure via which the expansions (9.64), (9.65), (9.67), (9.68) have been obtained; while the contents of this section can be made rigorous, the resulting proofs are long, elementary, and add little to the subject of this paper.

We start with introducing the notation

$$a_k = \frac{2k(k+1) - 1}{(2k+3)(2k-1)} \cdot c^2, \quad (9.40)$$

$$b_k = \frac{(k+2)(k+1)}{(2k+3)\sqrt{(2k+1)(2k+5)}} \cdot c^2, \quad (9.41)$$

so that the symmetric tridiagonal matrices $A^{\mu,\nu}$ assume the form

$$A_{k,k}^{\mu,\nu} = k(k+1) + a_k, \quad (9.42)$$

$$A_{k,k+2}^{\mu,\nu} = b_k, \quad (9.43)$$

$$A_{k+2,k}^{\mu,\nu} = b_k \quad (9.44)$$

(in Figure 1, $A^{(n-2),(n+2)}$ is depicted). Denoting by $A^{\mu,\nu} = \{(A^{\mu,\nu})_{i,j}\}$ the square submatrix of either A_{even} or A_{odd} (see Observation 5 above) consisting of elements $A_{i,j}$ such that $\mu \leq i, j \leq \nu$, we start with the following observation.

Observation 5. For a fixed c and any $0 < \mu < \nu$, the matrix $A^{\mu,\nu}$ is diagonally dominant. Moreover, for sufficiently large μ

$$A_{k+1,k+1}^{\mu,\nu} - A_{k,k}^{\mu,\nu} > |A_{k+1,k}^{\mu,\nu}| + |A_{k,k+1}^{\mu,\nu}| \quad (9.45)$$

for all $\mu < k < \nu - 1$. Thus, for any sufficiently large μ ,

$$\lambda_k(A^{\mu,\nu}) \sim 2k \cdot (2k + 1), \quad (9.46)$$

with $\lambda_k(A^{\mu,\nu})$ denoting the eigenvalue number k of the matrix $A^{\mu,\nu}$.

Consider now the submatrix $A^{(n-k),(n+k)}$ of the matrix A_{even} (see Figure 1), where the case of $k = 2$ is depicted). Obviously, for large n , it has an eigenvalue close to $\chi^0 = n(n + 1)$, with the corresponding eigenvector close to

$$x^0 = (\overbrace{0, 0, \dots, 0}^k, 1, \overbrace{0, \dots, 0}^k); \quad (9.47)$$

we would like to find an improved approximation to the eigenvalue of $A^{(n-k),(n+k)}$ closest to $n(n + 1)$, and to the corresponding eigenvector. Employing the standard inverse power method with a shift, we form the matrix

$$B^{(n-k),(n+k)} = A^{(n-k),(n+k)} - \chi^0 \cdot I, \quad (9.48)$$

shown in Figure 2 (with I denoting the identity), and evaluate the sequence of vectors

$$\left(B^{(n-k),(n+k)}\right)^{-1}(x_0), \left(B^{(n-k),(n+k)}\right)^{-2}(x_0), \left(B^{(n-k),(n+k)}\right)^{-3}(x_0), \dots \quad (9.49)$$

On inspection of the formulae (9.40)-(9.42), it is obvious that every coordinate of every element in the sequence (9.49) is a rational function of c, n . Now, constructing the sequence in the symbolic form (we used Mathematica), one can decompose it into a power series with respect to c, n ; the results of such expansion are presented in (9.64), (9.65), (9.67), (9.68) below.

9.9 Connections Between $\psi_m(1)$ and λ_m for Large m

The principal purpose of this section is Theorem 6 below, providing an exact expression for the eigenvalues λ_m^c as functions of c, m , and $\psi_m^c(1)$. Subsequently, asymptotic expansions for $\psi_m^c(1)$ are substituted into (9.58) provided by Theorem 6 to obtain asymptotic expansions for λ_m^c .

We start with a simple lemma describing the behavior of λ_m^c for small c .

Lemma 3. For any $m = 0, 1, 2, \dots$,

$$\lim_{c \rightarrow 0} \lambda_m^c \cdot \frac{(m!)^2 \Gamma(m + \frac{3}{2})}{i^m c^m \sqrt{\pi} (2m)!} = 1. \quad (9.50)$$

Proof. Differentiating (9.23) m times with respect to x and evaluating the result at $x = 0$, we obtain

$$\lambda_m^c \psi_m^{(m)}(0) = (ic)^m \int_{-1}^1 t^m \psi_m(t) dt, \quad (9.51)$$

or

$$\frac{\lambda_m^c \psi_m^{(m)}(0)}{i^m c^m} = \int_{-1}^1 t^m \psi_m(t) dt; \quad (9.52)$$

therefore,

$$\lim_{c \rightarrow 0} \frac{\lambda_m^c \psi_m^{(m)}(0)}{i^m c^m} = \lim_{c \rightarrow 0} \int_{-1}^1 t^m \psi_m(t) dt. \quad (9.53)$$

Now, using the combination of (9.33), (9.8), (9.9), (9.10), and (9.13), we rewrite (9.53) as

$$\lim_{c \rightarrow 0} \frac{\lambda_m^c (2m)! \sqrt{(m+1/2)}}{i^m c^m 2^m m!} = \lim_{c \rightarrow 0} \sqrt{(m+1/2)} \int_{-1}^1 t^m P_m(t) dt, \quad (9.54)$$

or

$$\lim_{c \rightarrow 0} \frac{\lambda_m^c (2m)!}{i^m c^m 2^m m!} = \frac{\sqrt{(\pi)} m!}{2^m \Gamma(m + \frac{3}{2})}, \quad (9.55)$$

or, finally,

$$\lim_{c \rightarrow 0} \frac{\lambda_m^c}{c^m} = \frac{\sqrt{(\pi)} (2m)! i^m}{2^m \Gamma(m + \frac{3}{2}) (m!)^2}. \quad (9.56)$$

The following theorem is one of principal results of this paper. It provides an explicit expression for λ_m^c in terms of $\psi_m(1)$.

Theorem 6. For any positive real c and integer $m \geq 0$,

$$\lambda_m^c = \frac{i^m \sqrt{\pi} c^m (m!)^2}{(2m)! \Gamma(m + 3/2)} \cdot e^{F(c)}, \quad (9.57)$$

where $F(c)$ is given by the formula

$$F(c) = \int_0^c \left(\frac{2(\psi_m^T(1))^2 - 1}{2\tau} - \frac{m}{\tau} \right) d\tau. \quad (9.58)$$

Proof. Suppose that c_0, c are two positive real numbers such that $0 < c_0 < c$. Integrating (9.28) from c_0 to c , we obtain

$$\log(\lambda_m^c) = \log(\lambda_m^{c_0}) + \int_{c_0}^c \frac{1}{2t} (2(\psi_m^t(1))^2 - 1) dt. \quad (9.59)$$

Due to (9.33),

$$\frac{1}{2t} (2(\psi_m^t(1))^2 - 1) = \frac{m}{t} + p(t, m), \quad (9.60)$$

with p a smooth function of t (see (9.67) below for a more detailed analysis of p); substituting (9.60) into (9.59), we have

$$\log(\lambda_m^c) = \log(\lambda_m^{c_0}) + m \cdot (\log(c) - \log(c_0)) + \int_{c_0}^c p(t, m) dt \quad (9.61)$$

exponentiating (9.61) and using (9.50), we obtain

$$\lambda_m^c = \lambda_m^{c_0} \cdot c^m \cdot c_0^{-m} \cdot e^{\int_{c_0}^c p(t, m) dt}. \quad (9.62)$$

Remembering that p is a smooth function of t and using (9.50), we rewrite (9.62) as

$$\lambda_m^c = \lim_{c_0 \rightarrow 0} \left(\frac{\lambda_m^{c_0}}{c_0^{-m}} \right) \cdot c^m \cdot e^{\int_0^c p(t, m) dt} = \frac{i^m \sqrt{\pi} (m!)^2 c^m}{(2m)! \Gamma(m + 3/2)} \cdot e^{\int_0^c p(t, m) dt}, \quad (9.63)$$

which we combine with (9.60) to obtain (9.57), (9.58).

9.10 Formulae for PSWFs and Eigenvalues χ_m, λ_m

The procedure described in Section 9.8 has been implemented in Mathematica, and used to obtain asymptotic expansions for the eigenvalues χ_m^c and eigenfunctions ψ_m^c of the differential operator (9.1); the obtained asymptotic expansions were verified by substitution into (9.2), with the help of (9.7), (9.4), (9.13). The resulting expansions are listed in this section; several numerical examples illustrating their effectiveness can be found in Section 9.14 below.

9.11 Formulae for $\chi_m^c, \psi_m^c, \psi_m^c(1), \psi_m^c(0)$

Theorem 7. For all real $c > 0$ and integer $m > 0$,

$$\begin{aligned} \chi_m^c = & m(m+1) + \frac{c^2}{2} + \frac{c^2(4+c^2)}{32m^2} - \frac{c^2(4+c^2)}{32m^3} + \frac{c^2(28+13c^2)}{128m^4} - \\ & \frac{c^2(20+11c^2)}{64m^5} + \frac{c^2(3904+3936c^2+160c^4+5c^6)}{8192m^6} - \\ & \frac{c^2(5824+8416c^2+480c^4+15c^6)}{8192m^7} + c^2 \cdot O\left(\frac{c^8}{m^8}\right). \end{aligned} \quad (9.64)$$

Theorem 8. For all real $c > 0$, integer $m > 0$, and $x \in [-1, 1]$,

$$\begin{aligned}
\psi_m^c(x) = & \left(\frac{c^{14}}{1352914698240 m^7} \right) \cdot \overline{P_{m-14}}(x) + \left(\frac{c^{12}}{671088640 m^7} + \frac{c^{12}}{12079595520 m^6} \right) \cdot \overline{P_{m-12}}(x) \\
& + \left(\frac{c^{10} (153600 - c^4)}{193273528320 m^7} + \frac{5 c^{10}}{50331648 m^6} + \frac{c^{10}}{125829120 m^5} \right) \cdot \overline{P_{m-10}}(x) \\
& + \left(\frac{c^8 (262400 - 11 c^4)}{2013265920 m^7} + \frac{c^8 (55680 - c^4)}{2013265920 m^6} + \frac{c^8}{196608 m^5} + \frac{c^8}{1572864 m^4} \right) \cdot \overline{P_{m-8}}(x) \\
& + \left(\frac{c^6 (417955840 + 983040 c^2 - 46080 c^4 + c^8)}{64424509440 m^7} + \frac{c^6 (99072 - 13 c^4)}{50331648 m^6} \right. \\
& \left. + \frac{c^6 (15616 - c^4)}{25165824 m^5} + \frac{3 c^6}{16384 m^4} + \frac{c^6}{24576 m^3} \right) \cdot \overline{P_{m-6}}(x) + \\
& \left(\frac{c^4 (19169280 + 589824 c^2 + 8064 c^4 + c^8)}{201326592 m^7} + \frac{c^4 (27721728 + 393216 c^2 - 10240 c^4 + c^8)}{805306368 m^6} \right. \\
& \left. + \frac{c^4 (1920 - c^4)}{131072 m^5} + \frac{c^4 (2880 - c^4)}{393216 m^4} + \frac{c^4}{256 m^3} + \frac{c^4}{512 m^2} \right) \cdot \overline{P_{m-4}}(x) \\
& + \left(\frac{c^2 (3397386240 + 2415919104 c^2 + 21528576 c^4 - 1769472 c^6 - 43008 c^8 - c^{12})}{38654705664 m^7} \right. \\
& + \frac{c^2 (1523712 + 491520 c^2 + 29568 c^4 + c^8)}{25165824 m^6} + \frac{c^2 (540672 + 98304 c^2 - 4992 c^4 + c^8)}{12582912 m^5} \\
& + \frac{c^2 (512 - c^4)}{16384 m^4} + \frac{c^2 (256 - c^4)}{8192 m^3} + \frac{c^2}{32 m^2} + \frac{c^2}{16 m} \left. \right) \cdot \overline{P_{m-2}}(x) \\
& + \left(1 + \frac{31948800 c^4 + 589824 c^6 + 2688 c^8 + c^{12}}{201326592 m^7} - \frac{46301184 c^4 + 589824 c^6 - 1152 c^8 + c^{12}}{603979776 m^6} \right. \\
& + \frac{3840 c^4 - c^8}{131072 m^5} - \frac{4352 c^4 - c^8}{262144 m^4} + \frac{c^4}{256 m^3} - \frac{c^4}{256 m^2} \left. \right) \cdot \overline{P_m}(x) \\
& + \left(\frac{c^2 (-94296342528 - 11475615744 c^2 + 53968896 c^4 + 1769472 c^6 - 12288 c^8 + c^{12})}{38654705664 m^7} \right. \\
& + \frac{c^2 (28164096 + 1474560 c^2 - 66432 c^4 + 11 c^8)}{25165824 m^6} - \frac{c^2 (6832128 + 98304 c^2 - 17280 c^4 + c^8)}{12582912 m^5} \\
& - \frac{c^2 (-4608 + 7 c^4)}{16384 m^4} + \frac{c^2 (-1280 + c^4)}{8192 m^3} + \frac{3 c^2}{32 m^2} - \frac{c^2}{16 m} \left. \right) \cdot \overline{P_{m+2}}(x) \\
& + \left(\frac{-(c^4 (297271296 + 2359296 c^2 - 81152 c^4 + 5 c^8))}{402653184 m^7} + \right. \\
& \frac{c^4 (185008128 + 393216 c^2 - 61440 c^4 + c^8)}{805306368 m^6} - \frac{7 c^4 (4224 - c^4)}{393216 m^5} + \frac{c^4 (9792 - c^4)}{393216 m^4} \\
& \left. - \frac{c^4}{128 m^3} + \frac{c^4}{512 m^2} \right) \cdot \overline{P_{m+4}}(x) + \\
& \left(\frac{-c^6 (2053734400 + 983040 c^2 - 184320 c^4 + c^8)}{64424509440 m^7} + \frac{c^6 (367872 - 23 c^4)}{50331648 m^6} \right. \\
& - \frac{c^6 (40192 - c^4)}{25165824 m^5} + \frac{5 c^6}{16384 m^4} - \frac{c^6}{24576 m^3} \left. \right) \cdot \overline{P_{m+6}}(x) \\
& + \left(\frac{c^8 (-775680 + 17 c^4)}{2013265920 m^7} + \frac{c^8 (119680 - c^4)}{2013265920 m^6} - \frac{c^8}{131072 m^5} + \frac{c^8}{1572864 m^4} \right) \cdot \overline{P_{m+8}}(x) \\
& + \left(\frac{c^{10} (-291840 + c^4)}{193273528320 m^7} + \frac{7 c^{10}}{50331648 m^6} - \frac{c^{10}}{125829120 m^5} \right) \cdot \overline{P_{m+10}}(x)
\end{aligned} \tag{9.65}$$

The following two theorems provide asymptotic expansions for $\psi_m^c(1)$, $\psi_m^c(0)$. It should be observed that the expansions below differ from expansion (9.65), in that (9.65) is an expansion in powers of $\frac{c^2}{m}$, while the expansions (9.67), (9.68) below (as well as the expansion (9.64) above) are in powers of $\frac{c}{m}$. In numerical terms, it means that expansions (9.67), (9.68), (9.64) produce reasonable accuracy whenever $m \geq c$, while (9.65) is useless unless $c < m^2$, which tends to be a rather restrictive condition.

Theorem 9. For any $c > 0$ and integer $m \geq 0$,

$$|\psi_m^c(1)| < \sqrt{m + \frac{1}{2}}, \quad (9.66)$$

and

$$\begin{aligned} \psi_m^c(1) = & \sqrt{\frac{1}{2} + m} \left(1 - \frac{4c^2 + c^4}{4(1+2m)^4} - \frac{8c^2 + 5c^4}{(1+2m)^6} - \frac{3072c^2 + 6800c^4 + 496c^6 + 17c^8}{64(1+2m)^8} \right. \\ & - \frac{512c^2 + 4512c^4 + 1078c^6 + 55c^8}{2(1+2m)^{10}} \\ & - \frac{327680c^2 + 11668480c^4 + 6912192c^6 + 615944c^8 + 8844c^{10} + 107c^{12}}{256(1+2m)^{12}} \\ & + \frac{-393216c^2 - 56369152c^4 - 77568128c^6 - 12546920c^8 - 453880c^{10} - 7125c^{12}}{64(1+2m)^{14}} \\ & + \frac{-469762048c^2 - 270135197696c^4 - 847030386688c^6 - 247641926912c^8}{16384(1+2m)^{16}} \\ & + \frac{-16252261888c^{10} - 380878144c^{12} - 2115648c^{14} - 12573c^{16}}{16384(1+2m)^{16}} \\ & + \frac{-33554432c^2 - 77265371136c^4 - 548143177728c^6 - 287562215936c^8}{256(1+2m)^{18}} \\ & + \frac{-31109328352c^{10} - 1111543216c^{12} - 14134828c^{14} - 101755c^{16}}{256(1+2m)^{18}} \\ & + \frac{-38654705664c^2 - 356175128297472c^4 - 5699666788220928c^6}{65536(1+2m)^{20}} \\ & + \frac{-5342103064903680c^8 - 921930796096512c^{10} - 49298957839616c^{12}}{65536(1+2m)^{20}} \\ & + \frac{-1022214431488c^{14} - 10219338512c^{16} - 28971260c^{18} - 100327c^{20}}{65536(1+2m)^{20}} \\ & + O\left(\frac{1}{m^{22}}\right). \end{aligned} \quad (9.67)$$

Theorem 10. For any $c > 0$ and integer $m \geq 1$,

$$\begin{aligned}
\psi_m^c(0) = & \sqrt{\frac{2}{\pi}} \left(1 - \frac{2c^2+1}{16m^2} + \frac{2c^2+1}{16m^3} + \frac{8c^4-140c^2-3}{512m^4} \right. \\
& - \frac{103680c^4 - 1399680c^2 + 4601}{3317760m^5} \\
& - \frac{83980800c^6 - 976276800c^4 + 7865326800c^2 - 17805089}{10749542400m^6} \\
& + \frac{67184640c^6 - 557072640c^4 + 3430412748c^2 + 86309}{2866544640m^7} \\
& + \frac{1}{12383472844800m^8} (20785248000c^8 - 1986313881600c^6 \\
& + 4275336556800c^4 - 24430778197776c^2 - 1978643839) \\
& - \frac{1}{24766945689600m^9} (166281984000c^8 - 13181626368000c^6 \\
& + 13528331886240c^4 - 79476319733832c^2 - 996600013) \\
& - \frac{1}{99067782758400m^{10}} (82763078400c^{10} - 4876030224000c^8 \\
& + 244104975820800c^6 - 21627672187536c^4 - 1276961293959615c^2 \\
& + 1381481054) + \frac{1}{792542262067200m^{11}} \\
& (3310523136000c^{10} - 155133532800000c^8 + 6730329331837440c^6 \\
& + 149489524096664544c^4 - 25501720589940204c^2 + 4407192151) \Big) \\
& + O\left(\frac{c^{12}}{m^{12}}\right).
\end{aligned} \tag{9.68}$$

9.12 Formulae for λ_m^c

Theorem 6 expresses λ_m^c via c , m , and $\psi_m^c(1)$, and Theorem 9 provides an expansion of $\psi_m^c(1)$ into powers of $1/m$. Combining these two observations (and carrying out the elementary but voluminous manipulations), we readily obtain expressions for λ_m^c , given in Theorem 11, proved by substituting (9.67) into (9.58) and carrying out the integration.

Theorem 11. Suppose that $c > 0$ is a real number, and m is a sufficiently large integer. Then

$$|\lambda_m^c| < \frac{c^m \sqrt{\pi} (m!)^2}{(2m)! \Gamma(\frac{3}{2} + m)}, \tag{9.69}$$

and

$$|\lambda_m^c| = \frac{c^m \sqrt{\pi} (m!)^2}{(2m)! \Gamma(\frac{3}{2} + m)} \cdot e^q \tag{9.70}$$

with

$$\begin{aligned}
q = & -\frac{8c^2 + c^4}{16(1+2m)^3} - \frac{16c^2 + 5c^4}{4(1+2m)^5} \\
& - \frac{12288c^2 + 53952c^4 + 8512c^6 + 315c^8}{96(1+2m)^9} - \frac{+12288c^2 + 13536c^4 + 640c^6 + 15c^8}{512(1+2m)^7} \\
& - \frac{327680c^2 + 5824000c^4 + 2290560c^6 + 150990c^8 + 1656c^{10} + 15c^{12}}{512(1+2m)^{11}} \\
& - \frac{196608c^2 + 14082048c^4 + 12891648c^6 + 1554795c^8 + 44088c^{10} + 550c^{12}}{64(1+2m)^{13}} \quad (9.71) \\
& - \frac{384978886656c^{10} + 7344995840c^{12} + 33146880c^{14} + 154245c^{16}}{3932160(1+2m)^{15}} \\
& - \frac{56371445760c^2 + 16203707842560c^4 + 33835125309440c^6 + 7399764503040c^8}{3932160(1+2m)^{15}} \\
& + O\left(\frac{c^{20}}{m^{19}}\right).
\end{aligned}$$

Remark 1. It should be observed that the expansion (9.71) has a drawback as a tool for approximating λ_m^c for large values of c . Indeed, some of the terms in (9.71) are of the form c^{k+1}/m^k with $k = 3, 7, 11, 15$. In other words, in order to obtain a prescribed accuracy for large c , the values of m have to grow *faster* than the values of c . For example, in the case of (9.71), m must grow as $c^{20/19}$. While numerical examples in Section 9.14 show that the resulting approximations are acceptable for fairly large ratios of c/m and fairly large c , in certain situations such estimates are not sufficient. To some extent, we attempt to remedy this problem in the following section.

Remark 2. Looking at the expansions (9.64), (9.67), (9.68) (but not the expansions (9.65), (9.71)), one is tempted to think that each of them represents the first few terms of a convergent series. We conjecture that indeed, there exist expansions approximating χ_m , $\psi_m^c(1)$, $\psi_m^c(0)$ that are convergent for $c < m$, and for which (9.64), (9.67), (9.68) are the first 7, 20, and 11 terms, respectively. Such convergent expansions are being investigated.

9.13 Results Following From the WKB Analysis of the Equation (9.2)

While all expansions in Section 9.10 are in terms of powers of c/m , it is often desirable to have approximate formulae for various quantities associated with the PSWFs that are valid when (for example) m increases, but c stays proportional to m . In this section, we list several such estimates. Since the

results presented here are not the principal purpose of this paper, they are listed without proofs. The proofs will be published at a later date.

Theorem 12. Suppose that $c \geq 1$, and that $m \geq c$ is an integer. Then

$$\left| \frac{\frac{\psi_m(1)}{\psi_m(0)}}{\sqrt{(\pi/2) \cdot (\chi_m^c)^{1/4}}} - 1 \right| = O\left(\frac{1}{m^2}\right). \quad (9.72)$$

Theorem 13. Suppose that $c \geq 1$, and that $m \geq c$ is an integer. Then

$$\left| \frac{\chi_m^c}{c^2 \cdot \overline{E}(k \cdot \pi / (2 \cdot c))} - 1 \right| = O\left(\frac{1}{\sqrt{(m \cdot c)}}\right), \quad (9.73)$$

with $\overline{E}$ defined in (9.20), (9.21).

Theorem 14. Suppose that $c \geq 1$, and that $m \geq c$ is an integer. Then

$$\left| \frac{|\lambda_m^c|}{p_0(c, m)} - 1 \right| < \frac{1}{\sqrt{c \cdot m}}, \quad (9.74)$$

where $p_0(c, m)$ is defined by the formula

$$p_0(c, m) = \frac{e^{-\sqrt{\chi_m^c} \cdot (F(\sqrt{1-c^2/\chi_m^c}) - E(\sqrt{1-c^2/\chi_m^c}))}}{\sqrt{c} \cdot \sqrt{2 \cdot \pi}}, \quad (9.75)$$

and E, F are the elliptic integral defined in (9.16), (9.17).

The above theorem provides an estimate that is effective for arbitrarily small c/m , and gets tighter as c increases. When c/m is reasonably close to 1, a tighter estimate is provided by Theorem 15 below; however, the estimate provided by Theorem 15 *deteriorates* as c/m decreases.

Theorem 15. Suppose that $b \leq 1$, and that $c = m \cdot b$. Then

$$\lambda_m^c = p_0(c, m) \cdot p_1(c, m) \cdot \left(1 + O\left(\frac{1}{m}\right)\right), \quad (9.76)$$

where p_0 is defined in (9.75), $p_1(c, m)$ is defined by the formula

$$p_1(c, m) = e^{\frac{(c^2 \cdot F(\sqrt{1-c^2/\chi_m^c}) - \chi_m^c \cdot E(\sqrt{1-c^2/\chi_m^c}))}{2 \cdot (\chi_m^c - c^2) \cdot \sqrt{(\chi_m^c)}}} \quad (9.77)$$

and E, F are the elliptic integrals defined in (9.16), (9.17).

Theorem 16. Suppose that $b \leq 1$, and that $c = m \cdot b$. Then for each $k = 1, 2, \dots, m$,

$$\left| x_k - \sin \left(\tilde{G} \left(\frac{k \cdot \pi}{\sqrt{\chi_m}}, \frac{c}{\sqrt{\chi_m}} \right) \right) \right| = O \left(\frac{1}{m^2} \right), \quad (9.78)$$

where x_k denotes the k -th root of ψ_m^c , and $\tilde{G}$ is defined in (9.19).

9.14 Numerical Results

The approximate formulae of Sections 9.10, 9.13 were tested numerically. The calculations were performed in FORTRAN (the LINUX version from Lahey), in either double or extended precision, as needed. Tables 1 - 8 below illustrate the numerical behavior of some of the approximations listed in Sections 9.10, 9.13 above. The first two columns in each of the tables contain the bandlimit c and the degree m of the functions for which the calculation was performed. The third column contains the value of the parameter being approximated (calculated numerically with sufficient accuracy via the algorithm described in [15]), and the fourth column contains the approximation being tested. Finally, the fifth column contains the relative error of the approximation. More specifically,

- the parameter $\widetilde{\chi_m^c}$ in Table 1 is defined in (9.64) in Theorem 7;
- the parameter $\widetilde{\psi_m^c(1)}$ in Table 2 is defined in (9.66) in Theorem 9;
- the parameter $\widetilde{\psi_m^c(0)}$ in Table 3 is defined in (9.68) in Theorem 10;
- the parameter $\widetilde{\lambda_m^c}$ in Table 4 is defined in (9.70) in Theorem 11;
- the function p_0 in Table 7 is defined in (9.75) in Theorem 14;
- the function p_0 in Table 8 is defined in (9.75), and the function p_1 in Table 8 is defined in (9.77).

The final table (Table 9) illustrates the performance of the approximation (9.76) in the regime where the authors have encountered the need for such an approximation most frequently - when attempting to determine the smallest order m such that

$$|\lambda_m^c| \leq \varepsilon, \quad (9.79)$$

with c a prescribed band-limit, and ε a reasonably small real number (see [15]). Here, $\widetilde{\lambda_m^c}$ is defined by (9.76).

The following observations can be made from the tables below, and from the more extensive numerical experiments performed by us.

1. Approximations obtained via the inverse power method of Section 9.8 tend to be surprisingly accurate; all of them display rapid convergence as $c/m \rightarrow 0$. With the exception of (9.70) in Theorem 11, they produce at least three digits at $m \sim c$, and better than two digits at $m \sim 0.7c$. The estimate (9.70) produces almost 4 digits at $m \sim c$, but breaks down quickly once $m < c$; generally, this estimate should be used with care (see Remark 1).
2. Approximations obtained via the WKB analysis of the equation (9.2) are slowly convergent (which is to be expected from (9.72), (9.73), (9.74), (9.76)), but quite robust. All of them are accurate to a few percent at $m \sim c \sim 10$, and produce better than two-digit accuracy at $m \sim c \sim 100$. The estimate (9.76) is an exception, in that it produces almost 4-digit accuracy at $m \sim c \sim 10$ (even though its convergence is only first order).
3. The obtained expansions tend to be highly satisfactory as “ballpark” estimates, or when initial approximations are needed for some iterative processes (Newton, inverse power with shifts, etc.). The authors have already used some of them for this purpose.
4. In all cases, convergence rates predicted by the estimates (9.64), (9.67), (9.68), (9.70), (9.72), (9.73), (9.74), (9.76)) are confirmed numerically.
5. The expansions (9.64), (9.67), (9.68), (9.70) of Section 9.8 have been obtained via the inverse power method followed by the expansion of the result into a power series in c/m (with some additional analysis in the case of (9.70)). Thus, convergence rates of expansions in Section 9.8 are fairly high, and can be easily improved (in fact, some of them were obtained with higher orders, and truncated in order to save space and improve readability).
6. In contrast, the expansions (9.72), (9.73), (9.74), (9.76) of Section 9.13 have been obtained via a WKB analysis of equation (9.2), and converge slowly; as often happens in the WKB environment, higher order terms can be obtained, but tend to require fairly involved algebraic manipulation.
7. Table (9) indicates that whenever $|\lambda_m^c| \in [10^{-3}, 10^{-34}]$, the approximation (9.73) to λ_m^c is quite accurate, even for $c = 1$. Because of this relative universality, we found the approximation (9.76) to be convenient in many practical situations.

9.15 Conclusions

In this paper, we continue to develop asymptotic formulae for the approximation of certain Prolate Spheroidal Wave functions; in this sense, this paper can be viewed as a sequel to [15], [16]). We investigate the behavior of PSWFs in two regimes: when the ratio c/m decays, and when both c and m grow, but the ratio c/m stays bounded. Both the regions of validity and the accuracies of the presented expansions are illustrated with numerical examples.

While our results are restricted to PSWFs representing band-limited functions in one dimension, they are easily extended to PSWFs in two and three dimensions (see [10]), and to the discrete version of PSWFs (see [14]).

References

- [1] C.J. Bouwkamp. On Spheroidal Wave Functions of Order Zero. *J. Math. Phys.*, 26:79–92, 1947.
- [2] M.V. Fedoryuk, *Asymptotic Analysis of Linear Ordinary Differential Equations*, Springer, 1993.
- [3] W.H. Fuchs. On the Eigenvalues of an Integral Equation Arising in the Theory of Band-Limited Signals. *Journal of Mathematical Analysis and Applications*, 9:317–330, 1964.
- [4] F. Gantmacher and M. Krein. *Oscillation Matrices and Kernels and Small Oscillations of Mechanical Systems*. Gosudarstv. Izdat. Tehn.-Teor. Lit., 2nd edition, 1950. Russian.
- [5] I.S. Gradshteyn and I.M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, Inc., 1980.
- [6] S. Karlin and W.J. Studden. *Tchebycheff Systems with Applications In Analysis And Statistics*. John Wiley (Interscience), 1966.
- [7] H.J. Landau and H.O. Pollak. Prolate Spheroidal Wave Functions, Fourier analysis, and Uncertainty - II. *The Bell System Technical Journal*, pages 65–94, 1 1961.
- [8] H.J. Landau and H. Widom. Eigenvalue Distribution of Time and Frequency Limiting. *Journal of Mathematical Analysis and Applications*, 77:469–481, 1980.

- [9] P.M. Morse and H. Feshbach. *Methods of Theoretical Physics*. McGraw-Hill, New York, 1953.
- [10] D. Slepian. Prolate Spheroidal Wave Functions, Fourier analysis, and Uncertainty - IV: Extensions to many Dimensions, Generalized Prolate Spheroidal Wave Functions. *The Bell System Technical Journal*, 43:3009–3057, 1964.
- [11] D. Slepian. On Bandwidth. *Proc. of the IEEE*, 64:292–300, 1976.
- [12] D. Slepian and H.O. Pollak. Prolate Spheroidal Wave Functions, Fourier analysis, and Uncertainty - I. *The Bell System Technical Journal*, pages 43–63, 1 1961.
- [13] Slepian, D. (1965). Some Asymptotic Expansions for Prolate Spheroidal Wave Functions, *J. Math. and Physics*, 44, 99-140.
- [14] D. Slepian, *Prolate Spheroidal Wave Functions, Fourier Analysis, and Uncertainty - V: The Discrete Case*, The Bell System Technical Journal, May-June 1978.
- [15] H. Xiao and V. Rokhlin. Prolate Spheroidal Wave Functions, Quadrature and Interpolation. *Inverse Problems*, 17:805–838, 2001.
- [16] H. Xiao and V. Rokhlin. High-Frequency Asymptotic Expansions for Certain Prolate Spheroidal Wave Functions. *Journal of Fourier Analysis and Applications*, 9:577–598, 2003.

$(n-4)(n-3)$ $+a_{n-4}$	b_{n-4}			
b_{n-4}	$(n-2)(n-1)$ $+a_{n-2}$	b_n		
	b_n	$n(n+1)$ $+a_n$	b_{n+2}	
		b_{n+2}	$(n+2)(n+3)$ $+a_{n+2}$	b_{n+4}
			b_{n+4}	$(n+4)(n+5)$ $+a_{n+4}$

Figure 1: Matrix $A^{(n-2),(n+2)}$

$(n-4)(n-3)$ $+a_{n-4}$ $-n(n+1)$	b_{n-4}			
b_{n-4}	$(n-2)(n-1)$ $+a_{n-2}$ $-n(n+1)$	b_n		
	b_n	a_n	b_{n+2}	
		b_{n+2}	$(n+2)(n+3)$ $+a_{n+2}$ $-n(n+1)$	b_{n+4}
			b_{n+4}	$(n+4)(n+5)$ $+a_{n+4}$ $-n(n+1)$

Figure 2: Matrix $B^{(n-2),(n+2)}$

Table 1: Illustration to Theorem 7

c	m	χ_m^c	$\widetilde{\chi_m^c}$	$ \frac{\widetilde{\chi_m^c}}{\chi_m^c} - 1 $
10	10	0.1630966527E+03	0.1630714644E+03	0.15444E-03
10	20	0.4707790239E+03	0.4707789503E+03	0.15648E-06
100	200	0.4527786640E+05	0.4527786600E+05	0.86985E-08
200	200	0.6146952030E+05	0.6146804219E+05	0.24046E-04
1000	1000	0.1532864705E+07	0.1532827516E+07	0.24261E-04
1000	700	0.1061453125E+07	0.1059550913E+07	0.17921E-02
2000	2000	0.6129524565E+07	0.6129375490E+07	0.24321E-04
2000	1400	0.4244667882E+07	0.4237027811E+07	0.17999E-02
2000	2800	0.9906881564E+07	0.9906876723E+07	0.48867E-06
2000	4000	0.1803528048E+08	0.1803528034E+08	0.74676E-08
3000	3000	0.1378997985E+08	0.1378964417E+08	0.24342E-04

Table 2: Illustration to Theorem 9

c	m	$\psi_m^c(1)$	$\widetilde{\psi_m^c(1)}$	$ \frac{\widetilde{\psi_m^c(1)}}{\psi_m^c(1)} - 1 $
10	10	0.31902788747723779E+01	0.31905422086869083E+01	0.82543E-04
10	20	0.45234567755878853E+01	0.45234567756167842E+01	0.63887E-11
100	200	0.14146047761502584E+02	0.14146047761502849E+02	0.18696E-13
200	200	0.13924692587818798E+02	0.13924695713143016E+02	0.22444E-06
1000	1000	0.31101114869861699E+02	0.31101121823492145E+02	0.22358E-06
1000	700	0.23879885740948034E+02	0.23942265326233700E+02	0.26122E-02
2000	2000	0.43977324022989220E+02	0.43977333899706241E+02	0.22459E-06
2000	1400	0.33756823931676613E+02	0.33845812629714789E+02	0.26362E-02
2000	2800	0.52700849165636936E+02	0.52700849168898644E+02	0.61891E-10
2000	4000	0.63187511935751733E+02	0.63187511935752509E+02	0.12285E-13
3000	3000	0.53858430607666211E+02	0.53858442724682153E+02	0.22498E-06

Table 3: Illustration to Theorem 10

c	m	$\psi_m^c(0)$	$\widetilde{\psi_m^c(0)}$	$ \frac{\widetilde{\psi_m^c(0)}}{\psi_m^c(0)} - 1 $
10	10	0.71144703287367743E+00	0.71171898554143206E+00	0.38225E-03
10	20	0.77458052422641627E+00	0.77458058647749832E+00	0.80367E-07
100	200	0.77375279244056135E+00	0.77375275802711210E+00	0.44476E-07
200	200	0.70560112377749721E+00	0.70550546307480897E+00	0.13557E-03
1000	1000	0.70524498101496006E+00	0.70514631315036206E+00	0.13991E-03
1000	700	0.59360424996183111E+00	0.59332063131563892E+00	0.47779E-03
2000	2000	0.70519974860604934E+00	0.70510078690664959E+00	0.14033E-03
2000	1400	0.59311889866962533E+00	0.59311889866962533E+00	0.45793E-03
2000	2800	0.74950308377257293E+00	0.74950078056463195E+00	0.30730E-05
2000	4000	0.77364297728974388E+00	0.77364294095129980E+00	0.46971E-07
3000	3000	0.70518463549707760E+00	0.70508558000608774E+00	0.14047E-03

Table 4: Illustration to Theorem 11

c	m	λ_m^c	$\widetilde{\lambda_m^c}$	$ \frac{\widetilde{\lambda_m^c}}{\lambda_m^c} - 1 $
10	10	0.74448729094687059E-02	0.74460264240524210E-02	0.15492E-03
10	20	0.11487284026412298E-09	0.11487284032200906E-09	0.50391E-09
100	200	0.13980106499480028E-94	0.13980106499962346E-94	0.34500E-10
200	200	0.50013133845361460E-35	0.50014447244613024E-35	0.26260E-04
1000	1000	0.21812532717459833-172	0.21815376765049754-172	0.13037E-03
1000	700	0.12446479810779478E-21	0.15087142653799010E-21	.17503E+00
2000	2000	0.83881420007423372-344	0.83903370115080797-344	0.26161E-03
2000	1400	0.21530614133976734E-42	0.31719477252513466E-42	.32122E+00
2000	2800	0.95872406658488436-883	0.95872445211599086-883	0.40213E-06
2000	4000	0.27032262499046933-939	0.27032262504623675-939	0.20630E-09
3000	3000	0.37241503575221057-515	0.37256142385539540-515	0.39292E-03

Table 5: Illustration to Theorem 12

c	m	$\frac{\psi_m^c(1)}{\psi_m^c(0)}$	$\sqrt{(\pi/2) \cdot (\chi_m^c)^{1/4}}$	$ \frac{\frac{\psi_m^c(1)}{\psi_m^c(0)}}{\sqrt{(\pi/2) \cdot (\chi_m^c)^{1/4}}} - 1 $
10	10	0.4484211371E+01	0.4478898292E+01	0.11848E-02
10	20	0.5839879308E+01	0.5838001665E+01	0.32152E-03
100	200	0.1828238670E+02	0.1828232508E+02	0.33704E-05
200	200	0.1973451022E+02	0.1973444398E+02	0.33567E-05
1000	1000	0.4409973230E+02	0.4409972636E+02	0.13475 E-6
1000	700	0.4022863000E+02	0.4022862080E+02	0.22871 E-6
2000	2000	0.6236151404E+02	0.6236151194E+02	0.33701E-07
2000	1400	0.5688802967E+02	0.5688802642E+02	0.57200E-07
2000	2800	0.7031438603E+02	0.7031438479E+02	0.17711E-07
2000	4000	0.8167528665E+02	0.8167528596E+02	0.84680E-08
3000	3000	0.7637493487E+02	0.7637493373E+02	0.14981E-07

Table 6: Illustration to Theorem 13

c	m	χ_m^c	$c^2 \cdot \overline{E}(k \cdot \pi / (2 \cdot c))$	$ \frac{\chi_m^c}{c^2 \cdot \overline{E}(k \cdot \pi / (2 \cdot c))} - 1 $
10	10	0.16310E+03	0.15319E+03	-0.60742E-01
10	20	0.47078E+03	0.45078E+03	-0.42476E-01
10	40	0.16902E+04	0.16502E+04	-0.23668E-01
100	200	0.45278E+05	0.45078E+05	-0.44093E-02
200	200	0.61470E+05	0.61276E+05	-0.31497E-02
1000	1000	0.15329E+07	0.15319E+07	-0.63084E-03
1000	700	0.10615E+07	0.10609E+07	-0.53822E-03
2000	2000	0.61295E+07	0.61276E+07	-0.31548E-03
2000	1400	0.42447E+07	0.42435E+07	-0.26871E-03
2000	2800	0.99069E+07	0.99041E+07	-0.28031E-03
2000	4000	0.18035E+08	0.18031E+08	-0.22135E-03
3000	3000	0.13790E+08	0.13787E+08	-0.21033E-03
4000	4000	0.24514E+08	0.24510E+08	-0.15775E-03

Table 7: Illustration to Theorem 14

c	m	$ \lambda_m^c $	$p_0(c, m)$	$ \frac{ \lambda_m^c }{p_0(c, m)} - 1 $
10	10	0.74449E-02	076927E-02	0.33293E-01
10	20	0.11487E-09	0.11744E-09	0.22321E-01
10	40	0.31613E-31	0.32045E-31	0.13658E-01
100	200	0.13980E-94	0.14011E-94	0.22430E-02
200	200	0.50013E-35	0.50097E-35	0.16745E-02
1000	1000	0.21813E-172	0.21820E-172	0.33499E-03
1000	700	0.12446E-21	0.12451E-21	0.38410E-03
2000	2000	0.83881E-344	0.83895E-344	0.16750E-03
2000	1400	0.21531E-42	0.21535E-42	0.19205E-03
2000	2800	0.95872E-883	0.95886E-883	0.14022E-03
2000	4000	0.18428E-1877	0.18430E-1877	0.11221E-03
3000	3000	0.37242E-515	0.37246E-515	0.11167E-03
4000	4000	0.17536E-686	0.17538E-686	0.83752E-04
5000	5000	0.852857-858	0.852914E-858	0.67002E-04

Table 8: Illustration to Theorem 15

c	m	$ \lambda_m^c $	$p_0(c, m) \cdot p_1(c, m)$	$ \frac{\lambda_m^c}{p_0(c, m) \cdot p_1(c, m)} - 1 $
10	10	0.7444872909E-02	0.7446665216E-02	0.24074E-03
10	20	0.1148728403E-09	0.1150119590E-09	0.12111E-02
10	40	0.3161339122E-31	0.3167374120E-31	0.19090E-02
100	200	0.1398010650E-94	0.1398175798E-94	0.11813E-03
200	200	0.5001313385E-35	0.5001357815E-35	0.88837E-05
1000	1000	0.2181253272E-172	0.2181257109E-172	0.17591E-05
1000	700	0.1244647981E-21	0.1244648034E-21	0.42152E-07
2000	2000	0.8388142000E-344	0.8388149368E-344	0.87843E-06
2000	1400	0.2153061413E-42	0.2153061458E-42	0.20906E-07
2000	2800	0.9587240665E-883	0.9587269587E-883	0.30167E-05
2000	4000	0.1842893028E-1877	0.1842903879E-1877	0.58880E-05
3000	3000	0.3724150357E-515	0.3724152537E-515	0.58536E-06
4000	4000	0.1753670459E-686	0.1753671228E-686	0.43892E-06
5000	5000	0.8528572778E-858	0.8528575773E-858	0.35108E-06

Table 9: Detailed test of approximation (9.76)

c	ε	m	$ \lambda_m^c $	$\widetilde{\lambda}_m^c$	$ \frac{\widetilde{\lambda}_m^c}{\lambda_m^c} - 1 $
1	10^{-3}	4	0.48332632960755855E-03	0.49253048415110964E-03	0.19043E-01
10	10^{-3}	12	0.36617050651533914E-03	0.36632358110157685E-03	0.41804E-03
100	10^{-3}	71	0.55575710332503292E-03	0.55575881296253393E-03	0.30762E-05
1000	10^{-3}	645	0.55106315217648969E-03	0.55107646952172820E-03	0.24167E-04
4000	10^{-3}	2554	0.90212772934407403E-03	0.90236151778056827E-03	0.25915E-03
1	10^{-7}	7	0.36792394641891258E-07	0.37488869176272692E-07	0.18930E-01
10	10^{-7}	17	0.49874818613646773E-07	0.49922485044638009E-07	0.95572E-03
100	10^{-7}	80	0.58925058618798641E-07	0.58925270851642587E-07	0.36017E-05
1000	10^{-7}	658	0.79048001359994678E-07	0.79048001648746860E-07	0.36529E-08
4000	10^{-7}	2571	0.61423452268446200E-07	0.61423452271823360E-07	0.54982E-10
1	10^{-16}	13	0.74952326257500369E-17	0.73796615083410797E-17	0.15661E-01
10	10^{-16}	27	0.15806109360487387E-16	0.15831555036994785E-16	0.16099E-02
100	10^{-16}	97	0.54155657267971079E-16	0.54156484982260280E-16	0.15284E-04
1000	10^{-16}	685	0.49840375608637398E-16	0.49840376752547266E-16	0.22951E-07
4000	10^{-16}	2603	0.76060266618255159E-16	0.76060266644868355E-16	0.34990E-09
1	10^{-34}	23	0.17086514327818415E-35	0.17289382300788054E-35	0.11873E-01
10	10^{-34}	43	0.66754620912146111E-35	0.66883769710725253E-35	0.19347E-02
100	10^{-34}	125	0.61670180107155583E-34	0.61672901821270542E-34	0.44133E-04
1000	10^{-34}	731	0.46579482948604880E-34	0.46579487739909906E-34	0.10286E-06
4000	10^{-34}	2661	0.55635248845674676E-34	0.55635248941472294E-34	0.17219E-08

10 A Wideband Fast Multipole Method for the Helmholtz Equation in Three Dimensions

10.1 Introduction

One of the standard approaches to the solution of scattering problems is to convert them into integral equations, which are then discretized using appropriate quadrature formulae. This usually leads to large-scale systems of linear algebraic equations, which are in turn solved via appropriately chosen iterative schemes (such as GMRES). Most iterative schemes for the solution of linear systems of this type require the application of the matrix of the system to a sequence of recursively generated vectors. Applying a dense matrix to a vector is an order N^2 procedure, where N is the dimensionality of the matrix, which in this case is equal to the number of nodes in the discretization of the domain of the integral equation (or to a small multiple of that number of nodes). As a result, the whole process is at least of order N^2 , which is prohibitive for many large-scale problems.

During the last 30 years or so, a number of algorithms have been constructed for the rapid application to arbitrary vectors of the matrices resulting from the discretization of integral equations of scattering theory. Historically, the first group of algorithms for this purpose were the so-called k -space methods, which take advantage of the fact that the free-space Green's function for the Helmholtz equation is translation invariant, and use Fast Fourier Transforms (see, for example, [6]). This approach can be quite efficient when applied to volume integrals, but usually requires order $N^{3/2} \cdot \log N$ operations when used for the solution of boundary integral equations in three dimensions. Its performance deteriorates in environments where some parts of the domain of the integral equation must be discretized at a higher resolution than other parts.

Another class of techniques is known as Fast Multipole Methods (FMMs). Algorithms in this class construct a hierarchical subdivision of the domain of the integral equation, in which it is recursively divided into smaller and smaller regions; for each region in the hierarchy, they use a "far field expansion" (in the original FMM for the Laplace equation, a multipole expansion; for low-frequency scattering problems, a partial-wave expansion) to represent the potential of that region, on regions distant from it. Since these expansions are used only on distant regions, arbitrary specified precision can be achieved (though the cost of the calculation grows somewhat when the required accuracy is increased). The CPU time requirements of the FMMs are proportional to N in the low-frequency regime, and to $N \cdot \log N$

for high-frequency problems involving boundary integrals; the latter estimate becomes $O(N)$ for volume integral equations. Fast Multipole Methods can easily be designed to be “adaptive”, that is, to subdivide the domain of the integral equation more finely in regions where the discretization contains more nodes; the performance of the FMM is only weakly affected (and often improved) in such environments.

While the algorithms known by the term “FMM” share the same basic computational structure, the expansions they use are based on either of two very different principles. For the Laplace equation, or for the Helmholtz equation in the low frequency regime, the principle used is that large submatrices of the matrix to be applied are of low rank (to high but finite precision); this permits a wide variety of expansions to be used successfully in these environments, as in, for instance, “fast” methods based on wavelets and similar structures (see, for example, [3], [2], [8]).

In the high frequency regime, the ranks of submatrices tend to be proportional to their sizes, and rank considerations can not be used to construct asymptotically fast algorithms. In this environment, existing fast schemes are based on a somewhat more subtle mathematical apparatus. Specifically, it turns out that the diagonal forms for the translation operators for the Helmholtz equation (see Section 10.12 below) are available analytically, and that on every level of subdivision *all translation operators are diagonalized by the same unitary operator* (see [27]). This observation leads to algorithms of order $N \cdot \log N$ for the application of discretized operators of scattering theory to arbitrary vectors, but not to algorithms of order N .

Each of these two types of expansions fails in some way outside its preferred regime: attempts to use rank-based approaches in the high-frequency regime result in algorithms whose CPU time requirements are proportional to N^2 ; attempts to use diagonal forms of translation operators at low frequencies result in numerically unstable schemes. In each case, the difficulty is fundamental, and can not be removed by simple expedients such as scaling, etc. Thus, there exist problems (e.g. scattering from an object many wavelengths in size which has significant subwavelength structure) for which neither of the two approaches performs well.

Fortunately, it is possible to construct “hybrid” schemes, conducting all calculations via partial-wave expansions on the subwavelength levels of subdivision, and transitioning to the diagonal form once the groups are of such size that the diagonal forms are stable. In this paper, we present such a procedure. More specifically, we describe an algorithm for the rapid

evaluation of expressions of the form

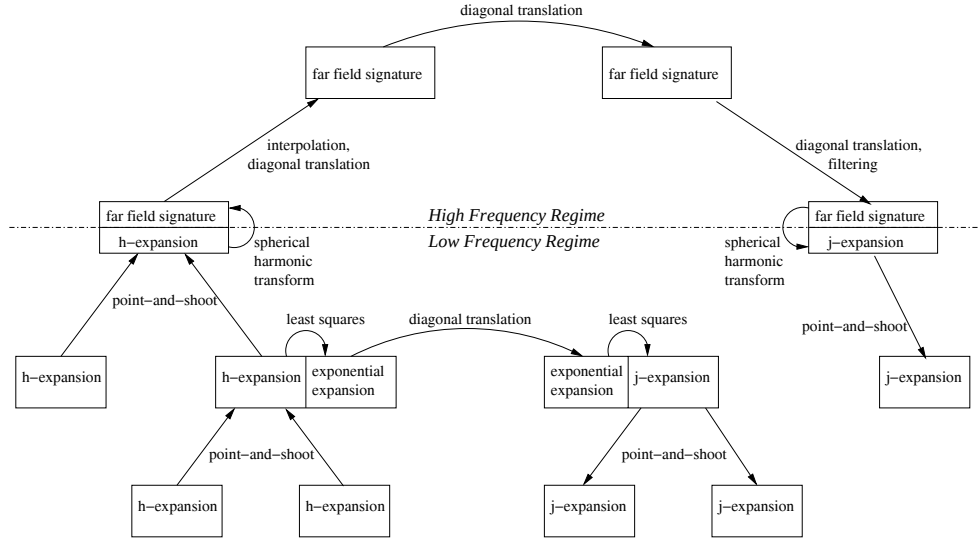
$$\Phi_k = \sum_{j=1, j \neq k}^N s_j \cdot h_0(\omega \cdot \|x_k - x_j\|), \quad (10.1)$$

or minor variants thereof, in which $\Phi_1, \dots, \Phi_N \in \mathbb{C}$ are to be calculated, given $x_1, \dots, x_N \in \mathbb{R}^3$, $s_1, \dots, s_N \in \mathbb{C}$, and $\omega \in \mathbb{C}$, and where h_0 denotes the spherical Hankel function of order zero ($h_0(z) = \frac{e^{iz}}{iz}$). In our description, we use the customary terminology of electromagnetics, referring to the formula

$$\Phi(x) = h_0(\omega \cdot \|x - x_0\|) \quad (10.2)$$

as the potential at the point x of a unit source of wavenumber ω located at the point x_0 . The potentials satisfy the law of superposition, so that, for instance, we refer to (10.1) as the potential at the points $x_1, \dots, x_N$ which is generated by sources $s_1, \dots, s_N$ located at those same respective points, all the sources being of the same wavenumber ω .

The algorithm presented here is a variant of the Fast Multipole Method (FMM); it is accurate and efficient for any value of ω whose real and imaginary parts are greater than zero, and we thus term it “wideband”. The basic computational structure of the wideband FMM is the same as that of previous variants of the FMM in three dimensions, and can be summarized as follows (for more detail, see Section 10.18). The algorithm finds the smallest cubical box which encloses all the points $\{x_j\}$, and then constructs a hierarchical subdivision of that box, in which it is divided into eight boxes of equal size (also cubes), each of which is likewise subdivided, the subdivision process continuing recursively until the lowest-level boxes have $O(1)$ points in them. The eight boxes into which a box is divided are termed its “children”, and it their “parent”. For each box, a far field expansion is produced, which represents the potential due to the sources on that box, at distances more than its own length away from it. The far field expansion for each childless (i.e. lowest-level) box is calculated from the sources on that box; the far field expansion for each parent box is calculated from the far field expansions of its children. These far field expansions are not evaluated directly (except in rare instances, having to do with the adaptive aspect of the algorithm), but rather translated into “local expansions”, which represent the potential inside a box due to sources distant from that box. Local expansions on parent boxes are evaluated by translating them into local expansions on their children; then the local expansion on each childless box is evaluated on the points $\{x_j\}$ in that box.



A small but representative fraction of the boxes used in a typical computation of the wide-band FMM. Boxes are placed in the figure according to their positions in the hierarchy, with parents on top of children. Inside each box is the type of expansion used for that box; on the left half of the figure, only far field expansions are listed, while on the right half, only local expansions are listed (although in the full computation, all boxes have both types of expansions). Where a box is divided into two, that indicates that two different types of expansions are used for that box. An arrow indicates a transformation performed between two expansions, and the text alongside it indicates the method used for that transformation. Exponential expansions are not fully represented in this figure: in the full computation, each box in the low-frequency regime has them; and not just one but six exponential expansions are used for each box.

Figure 10.1 illustrates the various expansions and methods for the translation and conversion of expansions used by the algorithm of this paper. At low frequencies, the principal expansions used are partial-wave expansions (Section 10.8), far field expansions of this type being referred to as “h-expansions” and local expansions as “j-expansions”. Translations of these expansions up and down in the hierarchy (for h-expansions, from child to parent; for j-expansions, from parent to child) are done using the “point-and-shoot” method (Section 10.11), in which the expansion is first rotated so that its z axis points in the direction in which it is to be translated, then translated, then rotated back. If the order of the expansion is denoted by p (the number of terms being p^2), then point-and-shoot is an $O(p^3)$ operation, whereas performing the translation via a single matrix multiplication would be $O(p^4)$. Translations from h-expansions to j-expansions are performed using exponential expansions (Section 10.16). This does not affect the asymptotic CPU time of the algorithm, but improves the constant substantially: conversion of far-field expansions into local expansions consumes most of the CPU time of the low-frequency portion of the algorithm (since each far field expansion might be translated to as many as $6^3 - 3^3 = 189$ local expansions, that being the maximum possible number of boxes on that level which are not adjacent to the source box, but whose parents are adjacent to its parents); since translation operators for exponential expansions are diagonal each of them can be applied in $O(p^2)$ time.

It has been empirically determined that the CPU time requirements of the scheme are minimized when high-frequency techniques are used wherever possible, which is whenever the box size exceeds a certain threshold (which depends on the desired accuracy of the computation, but is on the order of a wavelength). The high-frequency techniques are reviewed in Section 10.12; there, the far field expansion for a box (its “far field signature”) consists of the potential due to sources on that box, sampled on the surface of a sphere centered on the box, appropriately scaled, and taken in the limit as the radius of the sphere approaches infinity. The analytical machinery treats far field signatures as functions on the sphere; translation is performed by multiplying far-field signatures by other functions on the sphere. Numerically, far field signatures must also be resampled when they are translated; for downward translation, this amounts to filtering, and for upward translation, interpolation. Both interpolation and filtering are performed via a version of the algorithm introduced in [22] and summarized in Section 10.6 below.

The history of “fast” techniques of the type used in this paper goes back about 20 years; an excellent review can be found in [25]. An FMM for the Laplace equation in two dimensions was published in [18]. It has been

understood for a long time that trivial modifications convert the scheme of [18] into a viable FMM for the low-frequency Helmholtz equation; for algorithms of this type see [25] and references therein. In [21], the use of “intermediate” exponential expansions to accelerate the 2D Laplace FMM was proposed, and in [19], a three-dimensional version of that accelerated FMM was introduced. In [20], exponential expansions were worked out for the Helmholtz equation in three dimensions.

A high-frequency FMM for the Helmholtz equation in two dimensions was published in [26]; diagonal forms for translation operators in three dimensions are described in [27], and a single-level FMM in three dimensions is described in [9]. In [16], the complete theory of such diagonal forms is developed, and full-scale FMMs for the Helmholtz and Maxwell’s equations are constructed in [28], [33], [12], [13], [15].

In [5], an algorithm of a somewhat different type for the compression of matrices resulting from the discretization of integral equations of scattering theory is presented, both in the Helmholtz and Maxwell environments; the scheme is based on the observation that the free space Green’s functions for these equations are convolutions, and uses FFTs. An interesting development of this approach (under certain conditions, leading to remarkably efficient schemes) can be found in [32].

A detailed exposition of both (low-frequency and high-frequency) FMMs can be found in [7]. An outline of the wide-band FMM can also be found in [7]; a version of it based entirely on exponential expansions of the type described in Section 10.16 below was published in [14]. A scheme somewhat related to the one described here can be found in [23].

The structure of this paper is as follows. Section 10.2 contains the mathematical facts to be used in the paper that are either well-known or trivially follow from well-known facts. In Section 10.8, we discuss various properties of partial wave expansions that are relevant to the algorithm to be constructed, and introduce the concept of translation operators for partial wave expansions. Diagonal forms of translation operators in the high-frequency regime are studied in Section 10.12, and Section 10.16 is devoted to the diagonal forms of translation operators in the low-frequency regime (obtained via the so-called exponential expansions). The wideband FMM is described in Section 10.18, and the performance of the FMM is illustrated in Section 10.19 with a number of numerical examples. Finally, Section 10.20 contains very brief conclusions.

10.2 Mathematical Preliminaries

10.3 Symmetry

In describing various expansions and conversions between them, this paper makes use of a symmetry between far-field and local expansions, as follows. If we view (10.1) as a matrix-vector multiplication, and denote by A the matrix involved, then clearly A is symmetric; that is to say, the potential at any point x_1 due to a unit source at another point x_2 is equal to the potential at x_2 due to a unit source at x_1 . Suppose X_1 and X_2 are two subsets of the points $\{x_k\}$; denote by A_{12} the submatrix of A mapping sources at points in X_2 to potentials at points in X_1 , and by A_{21} the submatrix of A mapping sources at points in X_1 to potentials at points in X_2 . Since A is symmetric, $A_{12} = A_{21}^T$. If X_1 consists of the points on a box, and X_2 consists of the points for which that box's far field expansion is valid, then for any type of far field expansion for that box, there is a method of creating two matrices: a matrix C which creates the coefficients of the expansion, and a matrix E which evaluates the expansion, such that $\|A_{21} - EC\| < \varepsilon$, where ε is the accuracy of the expansion. Since A is symmetric, this also means that $\|A_{12} - C^T E^T\| < \varepsilon$; that is to say, for any type of far field expansion, there is a corresponding type of local expansion, such that the creation matrix for the far field expansion on any given box is the transpose of the evaluation matrix for the local expansion on that box, and the evaluation matrix for the far field expansion is the transpose of the creation matrix for the local expansion. It does not necessarily follow that the types of far field and local expansion used in any particular version of the FMM have this symmetry; but in almost all of the versions we are aware of, and in all the expansions used in this paper, this is in fact the case, if minor changes such as rescalings and sign changes of the coefficients of the expansions are neglected.

This symmetry also extends to conversions between two far field expansions, as compared to conversions between the two corresponding local expansions: in that case we have a far field creation matrix C , a far field evaluation matrix E for a different expansion (perhaps of a different type, or perhaps of the same type but for the parent box), and a conversion matrix B , such that $\|A_{21} - EBC\| < \varepsilon$. Then $\|A_{12} - C^T B^T E^T\| < \varepsilon$; that is to say, if we have a conversion matrix between two far field expansions, then the conversion matrix between the corresponding local expansions is its transpose.

Issues of numerical stability are likewise common to both types of expansion: if applying a sequence of matrices to a vector, one after another,

is a numerically stable process, then so is applying, in the reverse order, the sequence of transposes of those matrices.

The situation is slightly complicated by the presence of dipoles or of other more complicated sources: if those are used, the resulting analog of (10.1) is no longer symmetric. However, the algorithm still uses the same types of expansions in that situation — that is, although the creation matrices for far field expansions change, the evaluation matrices do not, nor do creation and evaluation matrices for local expansions. Thus translation matrices and conversion matrices between different types of expansions remain the same (and thus retain the above-described symmetry). Furthermore, if the desired result at each point were not the potential but an appropriate directional derivative of the potential, then the symmetry would be fully restored. Thus, even with this complication, the symmetry of far field and local expansions permits us, in many places in this paper, after treating far field expansions in full, to merely state that the local expansions are the corresponding ones, rather than repeating from a slightly different perspective their identical mathematics.

10.4 Spherical Harmonics

We use the usual spherical coordinates, in which the mapping from spherical coordinates (r, θ, ϕ) to Cartesian coordinates (x, y, z) is given by the formulae

$$x = r \sin \theta \cos \phi, \quad (10.3)$$

$$y = r \sin \theta \sin \phi, \quad (10.4)$$

$$z = r \cos \theta. \quad (10.5)$$

We denote by S^2 the two-dimensional sphere, each point $s \in S^2$ having coordinates $(\theta(s), \phi(s))$; for brevity, a point in spherical coordinates may be written (r, s) , meaning $(r, \theta(s), \phi(s))$.

A function $f : S^2 \rightarrow \mathbb{C}$ is referred to as a spherical harmonic of degree n if the function

$$r^n \cdot f(\theta, \phi) \quad (10.6)$$

satisfies the Laplace equation in $\mathbb{R}^3$. As is well known (see, for example, [11]), for any integer $n \geq 0$, there exist exactly $2n + 1$ linearly independent spherical harmonics of degree n , an orthonormal basis for which consists of the functions

$$Y_n^m(\theta, \phi) = \overline{P}_n^{|m|}(\cos \theta) e^{im\phi}, \quad (10.7)$$

for integer $m \in [-n, n]$, where $\overline{P}_n^m$ denotes the normalized associated Legendre function of degree n and order m , defined by the formula

$$\overline{P}_n^m(z) = \sqrt{\frac{2n+1}{4\pi}} \sqrt{\frac{(n-|m|)!}{(n+|m|)!}} \cdot P_n^m(z), \quad (10.8)$$

where P_n^m denotes the associated Legendre function of degree n and order m , which is given by Rodrigues' formula

$$P_n^m(x) = (-1)^m (1-x^2)^{(m/2)} \frac{d^m}{dx^m} P_n(x), \quad (10.9)$$

and where P_n denotes the Legendre polynomial of degree n . Any two spherical harmonics of different degrees are orthogonal, so the functions Y_n^m are all orthogonal to each other.

10.5 Integration of Spherical Harmonics

The grid of points used in the wideband FMM for sampling of functions on the sphere is the one introduced in [27]. It consists of the points (θ_k, ϕ_j) , for $k = 1, \dots, n$ and $j = 1, \dots, 2n$, where the points $\{\phi_j\}$ are equispaced nodes on the circle (that is, on $[0, 2\pi]$), and the points $\{\theta_k\}$ are given by the formula

$$\theta_k = \arccos t_k, \quad (10.10)$$

where $t_1, \dots, t_n \in [-1, 1]$ are the nodes of the n -point Gaussian quadrature on $[-1, 1]$. As is well known (see, for instance, [29]), the quadrature consisting of n equispaced nodes on the circle, with all quadrature weights equal to $2\pi/n$ (the trapezoidal rule) integrates exactly all functions of the form $e^{im\phi}$ with integer m such that $-n < m < n$, and the n -point Gaussian quadrature integrates polynomials of degree $2n-1$ exactly: denoting the weights of that quadrature by $w_1, \dots, w_n$, we have

$$\sum_{j=1}^n w_j t_j^m = \int_{-1}^1 t^m dt, \quad (10.11)$$

for any integer $m \in [0, 2n-1]$.

For integration on the sphere, the weight at each grid point (θ_k, ϕ_j) is taken to be $w_k \cdot \frac{2\pi}{n}$. This quadrature integrates exactly any spherical harmonic of degree less than $2n$: when applied to a function Y_n^m , if $m \neq 0$,

the integral on the sphere is zero, as is the result of applying the quadrature, while if $m = 0$, the integral of Y_n^m on the sphere is equal to

$$2\pi \int_0^\pi \overline{P}_n^{|m|}(\cos \theta) \sin \theta \, d\theta. \quad (10.12)$$

The substitution $t = \cos \theta$ converts (10.12) into the integral of a polynomial (due to (10.9), from which can be seen that the product of any two Legendre functions of the same order is a polynomial); it is thus integrated exactly by the Gaussian quadrature in t .

Since the product of a spherical harmonic of degree n_1 and a spherical harmonic of degree n_2 is a sum of spherical harmonics of degree $n_1 + n_2$ or less, the above quadrature integrates exactly products of spherical harmonics, provided the sum of their degrees is less than $2n$.

10.6 Filtering and Interpolation of Spherical Harmonics

Translations using the high-frequency diagonal form (Section 10.12) involve filtering and interpolation of functions on the sphere of the form

$$f(\theta, \phi) = \sum_{n=0}^p \sum_{m=-n}^n \alpha_n^m Y_n^m(\theta, \phi). \quad (10.13)$$

In that context, the ideal filtering scheme is one in which the filtered function values, sampled on a coarser grid, are given by the same formula (10.13), with the same coefficients $\{\alpha_n^m\}$, except that coefficients with $n > p'$ are set to zero, for some $p' < p$. Likewise, the ideal interpolation scheme is one in which the interpolated function values are given by the same formula (10.13), with the same coefficients $\{\alpha_n^m\}$, only sampled on a finer grid. Either of these schemes can be accomplished by computing the coefficients $\{\alpha_n^m\}$, then using them to evaluate (10.13) on the desired grid of points.

The computation of the coefficients $\{\alpha_n^m\}$ from the values of f on S^2 is referred to as a spherical harmonic transform. Since the functions $\{Y_n^m\}$ are orthonormal on S^2 , with the complex conjugate of $Y_n^m(s)$ being $Y_n^{-m}(s)$, it can be performed via the formula

$$\alpha_n^m = \int_{S^2} f(s) Y_n^{-m}(s) \, ds. \quad (10.14)$$

As shown in Section 10.5, the quadrature defined in that section is exact for such integrals, provided that the number of nodes in the grid is sufficient. Using that quadrature with sufficiently many nodes, the number of nodes

in the θ direction being denoted by q , and the number of nodes in the ϕ direction being $2q$, (10.14) becomes

$$\alpha_n^m = \frac{2\pi}{q} \sum_{k=1}^q w_k \sum_{j=1}^{2q} f(\theta_k, \phi_j) Y_n^{-m}(\theta_k, \phi_j). \quad (10.15)$$

This computation can be split into in two steps: first the calculation of the sums

$$\beta_k^m = \frac{2\pi}{q} \sum_{j=1}^{2q} f(\theta_k, \phi_j) e^{-im\phi_j} \quad (10.16)$$

for all $k = 1, \dots, q$ and $m = -p, \dots, p$, where p denotes the maximum order n for which the coefficients $\{\alpha_n^m\}$ are to be computed, then the calculation of the sums

$$\alpha_n^m = \sum_{k=1}^p w_k \beta_k^m \bar{P}_n^m(\cos \theta_k), \quad (10.17)$$

for all $n = 0, \dots, p$ and $m = -n, \dots, n$. Since the nodes $\{\phi_j\}$ are equispaced on the circle, the first of these steps can be performed by q invocations of the Fast Fourier Transform (FFT), each taking $O(q \log q)$ CPU time.

If filtering or interpolation of the above-described type is desired, the reverse of the above process must then be performed — that is, the evaluation of (10.13) on a (possibly different) grid of points $(\tilde{\theta}_k, \tilde{\phi}_j)$, of size $\tilde{q} \times 2\tilde{q}$. This process can likewise be accelerated by splitting it into two steps: first the calculation of the sums

$$\tilde{\beta}_k^m = \sum_{n=|m|}^p \alpha_n^m \bar{P}_n^m(\cos \tilde{\theta}_k), \quad (10.18)$$

for all $k = 1, \dots, \tilde{q}$ and $m = -k, \dots, k$, then the calculation of the sums

$$f_k^j = \sum_{m=-n}^n \tilde{\beta}_k^m e^{im\phi_j}, \quad (10.19)$$

for all $k = 1, \dots, \tilde{q}$ and $j = 1, \dots, 2\tilde{q}$. The second of these steps can be performed by $\tilde{q}$ invocations of the FFT, provided that the nodes $\{\tilde{\phi}_j\}$ are equispaced on the circle.

In [22], it was observed that the second and third stages of this interpolation or filtering process (that is, the computations (10.17) and (10.18)),

combined, can be evaluated efficiently using the one-dimensional FMM, and that this, together with the use of the FFT for the other two stages, constitutes an efficient algorithm for filtering or interpolation on the sphere. In outline, that observation is as follows: the combination of (10.17) and (10.18), consists, for each $m = -p, \dots, p$, of multiplication by a matrix whose (i,j)'th entry is given by the formula

$$w_i \cdot \sum_{n=|m|}^p \overline{P}_n^m(\cos \theta_i) \overline{P}_n^m(\cos \tilde{\theta}_j); \quad (10.20)$$

substituting into that formula the Christoffel-Darboux formula for the associated Legendre functions, which is

$$(\tilde{\mu} - \mu) \sum_{n=|m|}^p \overline{P}_n^m(\mu) \overline{P}_n^m(\tilde{\mu}) = \varepsilon_{p+1}^m (\overline{P}_{p+1}^m(\tilde{\mu}) \overline{P}_p^m(\mu) - \overline{P}_p^m(\tilde{\mu}) \overline{P}_{p+1}^m(\mu)), \quad (10.21)$$

where

$$\varepsilon_n^m = \sqrt{(n^2 - m^2)/(4n^2 - 1)}, \quad (10.22)$$

shows that the matrix (10.20) can be applied to a vector using two invocations of the one-dimensional FMM, each of which computes sums of the form

$$f_i = \sum_{j=1}^N \frac{q_j}{x_i - y_j}. \quad (10.23)$$

The wideband FMM uses this algorithm both for filtering and interpolation on the sphere; the one-dimensional FMM used in it is a version of the algorithm described in [31].

10.7 Spherical Bessel and Hankel functions

In accordance with standard practice [1], we denote by j_n the spherical Bessel function of the first kind of order n , and by h_n the spherical Hankel function of order n . Functions of both types are ‘elementary’ functions; in particular,

$$j_0(z) = \frac{\sin z}{z}, \quad (10.24)$$

$$h_0(z) = \frac{e^{iz}}{iz}. \quad (10.25)$$

For large n , the asymptotic behavior of $j_n(z)$ and $h_n(z)$ is given by the formulae (see [1], 9.3.1, 9.3.2, 9.1.3)

$$j_n(z) \sim \frac{z^n \cdot e^{n+\frac{1}{2}}}{2 \cdot (2n+1)^{n+1}}, \quad (10.26)$$

$$h_n(z) \sim \frac{\sqrt{2} \cdot (2n+1)^n}{z^{n+1} \cdot e^{n+\frac{1}{2}}}. \quad (10.27)$$

For large z with $\text{Im}(z) \geq 0$, the asymptotic behavior of $j_n(z)$ and $h_n(z)$ is given by the formulae (see [1], 9.2.5, 9.2.7, 10.1.1)

$$\lim_{z \rightarrow \infty} z \cdot j_m(z) = \cos(z - (m+1)\pi/2) + O\left(\frac{e^{\text{Im}(z)}}{|z|^2}\right), \quad (10.28)$$

$$\lim_{z \rightarrow \infty} z \cdot h_m(z) = e^{i(z-(m+1)\pi/2)} + O\left(\frac{e^{-\text{Im}(z)}}{|z|^2}\right). \quad (10.29)$$

10.8 Partial-Wave Expansions

Classical partial-wave expansions are the primary type of expansion used by the low-frequency Helmholtz FMM. They are given by the following theorem (which can be found, for instance, in [24]), which expresses as a series the potential from sources inside a disk in $\mathbb{R}^3$ on points outside the disk:

Theorem 17. Suppose that the function $\Phi : \mathbb{R}^3 \rightarrow \mathbb{C}$ is given by the formula

$$\Phi(x) = \sum_{k=1}^N q_k h_0(\|x - x_k\|), \quad (10.30)$$

and that $\|x_k\| \leq a$ for each $k = 1, \dots, N$. Then, denoting the spherical coordinates of the point x by (r, θ, ϕ) , and the spherical coordinates of each point x_k by (r_k, θ_k, ϕ_k) ,

$$\Phi(x) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \alpha_n^m Y_n^m(\theta, \phi) h_n(\omega r) \quad (10.31)$$

for any point x such that $r > a$, where

$$\alpha_n^m = \sum_{k=1}^N q_k Y_n^{-m}(\theta_k, \phi_k) j_n(\omega r_k). \quad (10.32)$$

Expansions of the form (10.31) will be referred to as h-expansions. Since the potential at any point x due to a unit source at another point y is the same as the potential at y due to a unit source at x , the above expansion can also be used with the positions of sources and measurement points reversed:

Corollary 4. *Suppose the conditions of Theorem 17 hold, except that $x_k > a$ for $k = 1, \dots, N$. Then*

$$\Phi(x) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \beta_n^m Y_n^{-m}(\theta, \phi) j_n(\omega r) \quad (10.33)$$

for any point x such that $r < a$, where

$$\beta_n^m = \sum_{k=1}^N q_k Y_n^m(\theta_k, \phi_k) h_n(\omega r_k). \quad (10.34)$$

Expansions of the form (10.33) will be referred to as j-expansions. In the FMM, whenever an h- or a j-expansion is used for some box in the hierarchy of boxes, the center of that box is taken to be the origin of the system of spherical coordinates used in (10.31) - (10.32) or (10.33) - (10.34).

Since j-expansions are related to h-expansions in the manner described in Section 10.3 — that is, the creation and evaluation operators for each type of expansion are the transposes of the evaluation and creation operators, respectively, for the other type of expansion — the following error analysis for h-expansions applies equally well to j-expansions.

10.9 Truncation Error

Each h-expansion used in the 3D FMM represents the potential from sources inside a cubical box, at locations outside that box and its immediate neighbors. Denoting the length of the box by $2R$, and taking the center of the box as the origin of the system of spherical coordinates, any point (r_k, θ_k, ϕ_k) inside the box satisfies the inequality

$$r_k \leq \sqrt{3} R, \quad (10.35)$$

and any point (r, θ, ϕ) at which the h-expansion for that box must be accurate satisfies the inequality

$$r \geq 3R. \quad (10.36)$$

These constraints permit the h-expansion to be truncated to a manageable number of terms. Examination of (10.31,10.32) shows that the appropriate truncation point depends on the possible magnitudes of the products

$$Y_n^m(\theta, \phi) \cdot h_n(\omega r) \cdot Y_n^{-m}(\theta_k, \phi_k) \cdot j_n(\omega r_k) : \quad (10.37)$$

if, for some n and m , the magnitude of (10.37) is bounded by some number ε , and the accuracy budget can handle an error of size ε , then the term α_n^m can be omitted from the calculation. Since the functions $\{Y_n^m\}$ are orthonormal and smooth, and thus have maximum values roughly equal to one, they may be dropped from (10.37) with only a modest loss of accuracy, leaving it as:

$$h_n(\omega r) \cdot j_n(\omega r_k). \quad (10.38)$$

Since this is not a function of m (and for reasons of general convenience), the truncation used in the FMM is purely in n — that is to say, the terms retained in the series (10.31) are those for which $n = 0, \dots, p$, and $m = -n, \dots, n$, where p depends on the desired accuracy and on the size R of the box. That dependence is a matter of the numerical behavior of the spherical Bessel functions h_n and j_n .

For real arguments, that behavior can be summarized as follows. For $x > n$, both $h_n(x)$ and $j_n(x)$ are oscillatory; as $x \rightarrow \infty$, the period of their oscillation tends toward a constant, and the magnitude of their oscillation decays as $1/x$. For $x < n$, $h_n(x)$ is monotonically increasing, starting from large negative values; in the limit as $x \rightarrow 0$, it is proportional to x^{-n-1} . For $x < n$, $j_n(x)$ is also monotonically increasing, but starting from small positive values; in the limit as $x \rightarrow 0$, it is proportional to x^n . Thus the non-negative values of n can be partitioned into three regions, in each of which the behavior of the product (10.38), under the constraints (10.35,10.36), is of a different character:

- Region I: $0 \leq n \leq \sqrt{3}\omega R$; here, in (10.38), j_n might be evaluated in its oscillatory regime, and h_n always is; since both factors of the product can be on the order of one, coefficients α_n^m for all these values of n must generally be included in the computation to achieve even minimal accuracy. For a box W wavelengths on each side (the wavelength being equal to $\frac{2\pi}{\omega}$), there are approximately $5W$ integer values of n in this region.
- Region II: $\sqrt{3}\omega R < n \leq 3\omega R$; here, in (10.38), j_n is evaluated on an argument which is less than n , and h_n is evaluated in its oscillatory regime. Thus the value of (10.38) is bounded by $j_n(\sqrt{3}\omega R)$, which

decreases with increasing n . If the box is sufficiently large and/or the desired precision sufficiently small, this decrease permits the truncation point p to be placed in this region. For a box W wavelengths in size, there are approximately $4W$ integer values of n in this region.

- Region III: $3\omega R < n$; here, in (10.38), both j_n and h_n are evaluated on arguments which are less than n ; thus j_n continues its decrease, and h_n starts to increase in magnitude. If $3\omega R \ll n$, then the approximations (10.26), (10.27) apply, and thus the maximum possible value of (10.38) under the constraints (10.36), (10.35) can be approximated by

$$\left(\frac{\sqrt{3}}{3}\right)^n \frac{1}{3R \cdot (2n+1)}. \quad (10.39)$$

The accuracy of this asymptotic bound is not satisfactory for numerical use outside the asymptotic regime; accordingly, in the wideband FMM, we bound the significance of the coefficients α_n^m by evaluating the product $h_n(3\omega R) \cdot j_n(\sqrt{3}\omega R)$ numerically. Experimentally, this product decreases with increasing n throughout this region, not just in the asymptotic regime.

The calculation and use of coefficients α_n^m , for n in Region III, clearly is more delicate than in Regions I or II, since in Region III the coefficients α_n^m can have small values yet large significance, being multiplied by a large number ($h_n(x)$, for $x < n$) when used in the series (10.31). The most delicate case is that of the coefficients at the truncation point (those for which $n = p$), which at worst may be multiplied by $h_p(3\omega R)$ when the expansion is evaluated.

This difference in the scaling of coefficients renders the high-frequency diagonal form (described below in Section 10.12) unusable for boxes below a certain size (which depends on the desired accuracy, but is on the order of a wavelength); the details of the problem are described further in that section. However, the difference in scaling does not cause problems if the formulae (10.31), (10.32) are used directly, except that if the boxes are very small, it introduces a different type of numerical problem, namely overflow or underflow of the exponents of floating-point numbers. This can be avoided by rescaling expansions: replacing $h_n(x)$ with $x^n h_n(x)$, and $j_n(x)$ with $j_n(x)/x^n$. This rescaling must be carried through the entire analysis; but as that is both straightforward and tedious, we omit it in our descriptions in this paper.

The above count of terms in Regions I and II, combined with (10.39), yields the result that the order p of partial-wave expansion needed to achieve

a given accuracy ε is $O(W + \log \varepsilon)$, where W is the size of the box in wavelengths. (The number of terms in a partial-wave expansion of order p is $(p + 1)^2$).

In the above, only the case of real wavenumber ω has been considered. In the case of a complex wavenumber whose imaginary part is positive (yielding waves that decrease exponentially away from sources), it suffices to apply the above truncation criterion to the real part of the wavenumber. For wavenumbers with large imaginary parts, this is, however, not an optimal truncation; further savings are possible. If the wavenumber is purely imaginary, the method of [17] should be used, as it is optimized for that case.

10.10 Rotation

By “rotating” an h-expansion, we refer to the following task: given an expansion of the form (10.31) relative to one system of spherical coordinates, convert it into an equivalent expansion of the same form relative to another system of spherical coordinates which shares the same origin. Denoting the coordinates in the first system by (r, θ_1, ϕ_1) and in the second by (r, θ_2, ϕ_2) (since they share the same origin, the radius is the same in both systems), denoting the coefficients of the expansion in the first system by $\{\alpha_n^m\}$, and denoting the order of that expansion by p , the task is to find coefficients $\{\beta_n^m\}$ such that

$$\sum_{n=0}^p \sum_{m=-n}^n \alpha_n^m Y_n^m(\theta_1, \phi_1) h_n(\omega r) = \sum_{n=0}^p \sum_{m=-n}^n \beta_n^m Y_n^m(\theta_2, \phi_2) h_n(\omega r). \quad (10.40)$$

Since this is to be true for any r , it must be the case that

$$\sum_{m=-n}^n \alpha_n^m Y_n^m(\theta_1, \phi_1) = \sum_{m=-n}^n \beta_n^m Y_n^m(\theta_2, \phi_2), \quad (10.41)$$

for each $n = 0, \dots, p$. Each of these $p + 1$ subproblems amounts to converting a spherical harmonic of degree n from one orthonormal basis to another, which is an exact operation (thus justifying our above demand for equality between the two expansions). The wideband FMM uses only a small number of angles of rotation, and uses partial-wave expansions only in the low-frequency regime, where they have limited numbers of terms; thus rotation matrices to accomplish this task can easily be precomputed and stored. Various methods of computing rotation matrices are given in [4]; methods sufficient for an FMM are summarized in [17]. However, here, we need only

use the brute-force numerical procedure of evaluating each function Y_n^m , rotated as desired, at suitable grid points (such as those defined in Section 10.5), then performing a spherical harmonic transform (see Section 10.6) to convert those values at grid points to spherical harmonic expansion coefficients: this yields the m 'th column of the n 'th rotation matrix. Using precomputed rotation matrices, a rotation takes $O(p^3)$ CPU time.

Rotation matrices for j -expansions are identical to those for h -expansions, as is evident either from symmetry considerations (as per Section 10.3) or by noting that the function j_n may be substituted for h_n in the above derivation without changing its result.

10.11 Translation

One of the basic building blocks of the FMM is a procedure for the translation of a far field expansion for a child box into a far field expansion for its parent. For partial-wave expansions, due to the ease of rotations (see the preceding section), we need only examine translations along the z axis: any other translation can be performed by rotating the expansion so that its z axis points along the direction of the desired translation (that is, from the center of the child box to the center of the parent box), translating it along the z axis, then rotating it back — a procedure referred to as “point-and-shoot”. Since each of these three steps can be performed in $O(p^3)$ time (as shown above for rotations, and below for translation in z), this is more economical than performing the whole translation via the application of a single dense matrix, which would take $O(p^4)$ time.

Formulae for the translation of h -expansions along the z axis can be found in Chapter 5 of [7], and in [17] (which gives formulae for the case of a purely imaginary wavenumber; to use them in the present context, the wavenumber must be divided by i). They are given in the following theorem:

Theorem 18. Suppose a point c_1 lies on the positive z axis at a distance ρ from the origin, and suppose that the potential Φ , at all points whose distance from c_1 is greater than r_1 , is given by an expansion of the form (10.31) centered at c_1 , whose coefficients are $\{\alpha_n^m\}$. Then at all points whose distance from the origin is greater than $r_1 + \rho$, Φ is given by an expansion of the same form centered at the origin, whose coefficients $\{\beta_n^m\}$ are given by the formula

$$\beta_n^m = \sum_{n'=m}^{\infty} C_{n,n'}^m \alpha_{n'}^m, \quad (10.42)$$

where

$$C_{n,n'}^m = \sum_{k=m}^{\min(n,n')} \left(\frac{1}{2}\right) (-1)^{n'+k} (2n' + 1) \cdot \frac{(n' - m)!(n + m)!(2k)!i^{-(n'+n)}(-\omega\rho)^{-k}j_n(\omega\rho)}{(k + m)!(k - m)!(n' - k)!(n - k)!k!} \quad (10.43)$$

In the wideband FMM, matrices containing the coefficients $\{C_{n,n'}^m\}$ are precomputed; since only one such set of coefficients is needed at each level of the hierarchy, and since they are only computed in the low-frequency regime, the time for this precomputation is negligible. Working with these precomputed matrices, the CPU time taken to perform a translation via (10.42) is $O(p^3)$.

Due to the symmetry discussed in Section 10.3, translation of a j-expansion from parent to child along the z axis is the transpose of the above translation operation; thus a similar point-and-shoot procedure is used for it, using the transpose of each of the precomputed matrices used in the above procedure, and applying them in reverse order.

As a remark, the fact that translations along the z axis do not mix together coefficients with different values of m (and thus that they are $O(p^3)$) can be seen without delving into the full proof of Theorem 18. Such translations involve two h-expansions, each relative to a different system of spherical coordinates, but with the coordinate ϕ of any given point being the same in both systems. In each expansion, the coefficients whose superscript is m are multiplied by $e^{im\phi}$, forming the m 'th Fourier component in ϕ of the potential. For the two expansions to yield the same potential, it is necessary and sufficient that their respective Fourier components be equal, which implies that the coefficients of each expansion with any given m must depend only on the coefficients of the other expansion for that m .

10.12 High-Frequency Diagonal Form

In [27], forms of far field and local expansion for the Helmholtz equation were introduced which have the property that translation operators (far field to far field, far field to local, and local to local) are diagonal. They are defined in terms of partial-wave expansions, as follows. The far field expansion for a given box, which is referred to as its “far field signature”, consists of the

function $f : S^2 \rightarrow \mathbb{C}$ given by the formula

$$f(s) = \sum_{n=0}^{\infty} \sum_{m=-n}^n (-i)^{n+1} \alpha_n^m Y_n^m(s), \quad (10.44)$$

where $\{\alpha_n^m\}$ are the coefficients of the h-expansion for that box. The local expansion for a given box (also referred to as a far field signature) consists of the function $g : S^2 \rightarrow \mathbb{C}$ such that

$$g(s) = \sum_{n=0}^{\infty} \sum_{m=-n}^n (-i)^{n+1} \beta_n^m Y_n^m(s), \quad (10.45)$$

where $\{\beta_n^m\}$ are the coefficients of the j-expansion for that box.

Since h-expansions and j-expansions are related to each other in the manner described in Section 10.3 (that is, the creation operator for the former is the transpose of the evaluation operator for the latter), far field signatures inherit, through (10.44) and (10.45), that same relationship.

From (10.44), the definition (10.31) of h-expansions, and the asymptotic behavior of the function h_n for large arguments (10.29), it follows that the far field signature f of any given box has the property that

$$f(s) = \lim_{r \rightarrow \infty} \Phi(r, s) \cdot r \cdot e^{-i\omega r}, \quad (10.46)$$

where $\Phi : \mathbb{R}^3 \rightarrow \mathbb{C}$ is the potential due to the h-expansion on that box (and thus, in the FMM, due to sources on that box), and where the system of spherical coordinates (r, s) has as its origin the center of that box.

Likewise, one may regard the function g for a given box, as given by (10.45), as a source distribution on a sphere centered on that box, which possesses the property that as the radius of that sphere is taken to infinity, the source distribution being appropriately scaled and adjusted in phase, the potential generated by it approaches the potential on the box.

10.13 Discretization

It is evident from (10.44), (10.45) that conversion of a partial-wave expansion to a far field signature, or vice versa, amounts to a spherical harmonic transform, forward or inverse (plus appropriate scaling by powers of i). Thus the grid points defined in Section 10.5 are suitable for sampling far field signatures; as shown in Section 10.6, if a $p \times 2p$ grid of that type is used, it allows for exact transformations from functions sampled on the grid to coefficients $\{\alpha_n^m\}$ of a partial-wave expansion of degree $p - 1$, and vice versa.

Numerically, since the basis functions $\{Y_n^m\}$ are orthonormal, these conversions are very well conditioned. However, in the case that the coefficients of the partial-wave expansion are scaled differently from each other, as they are for boxes less than roughly a wavelength in size, these conversions degrade or destroy the accuracy of the expansion, depending on how severe the difference in scaling is. As shown in Section 10.9, the worst case ill-conditioning is $h_p(3\omega R)$, where p is the number of terms in the h-expansion, and R is half the length of the box. Accordingly, the wideband FMM proceeds by calculating this value for each box size, and not using the high-frequency diagonal form for any boxes for which it exceeds the allowable loss of precision (as given by the ratio between the accuracy desired and the accuracy provided by the form of floating-point arithmetic in use). Since boxes at a given level in the hierarchy are all of the same size, this means that a cutoff line is drawn, above which the high-frequency form is used, and below which low-frequency techniques are used; the more precision is required, the higher in the hierarchy the line is. (See Table 10 for experimentally determined cutoff lines for the accuracies 10^{-3} , 10^{-6} , and 10^{-9}).

10.14 Far Field to Far Field and Local to Local Translations

The following theorem is proven in [27]:

Theorem 19. Suppose $f_1 : S^2 \rightarrow \mathbb{C}$ and $f_2 : S^2 \rightarrow \mathbb{C}$ are two far field signatures for the same potential, centered at points c_1 and c_2 , respectively. Then

$$f_2(s) = f_1(s) \cdot e^{i\omega \cdot (c_2 - c_1, E(s))}, \quad (10.47)$$

for all $s \in S^2$, where $E(s)$ denotes the unit vector in the direction s .

Theorem 19 provides a way to translate far field signatures upwards in the hierarchy of boxes. Numerically, both the function f_1 (the child's expansion) and the function f_2 (the child's expansion recentered on the parent box) are discretized on grids of the type defined in Section 10.5, with the grid used for f_2 having roughly twice as many points in each direction. Upwards translation thus consists of interpolating f_1 to the finer grid, then performing the operation (10.47) at each grid point. Due to the relation (10.44), the type of interpolation defined in Section 10.6 is ideal for this purpose.

Theorem 19 also provides a way to generate far field signatures for the potential generated by a set of point sources, without going through partial-wave expansions: each point source has a far field signature, centered on

itself, which for simple sources (of the type (10.2)) is a constant; such far field signatures, one for each point in the box, can each be translated to the center of the box via (10.47), then added together. (Formulae for this, for both simple sources and dipoles, can be found in [27]). This can be used to generate far field signatures for boxes whose size is above the high-frequency cutoff but which contain too few source points to be worth subdividing (although as implemented, the wideband FMM simply subdivides those boxes anyway).

For the reasons described in Section 10.3, each local-to-local translation operator is the transpose of the far-field-to-far-field translation operator for the same geometry. Considered as operators on functions on the sphere, both are diagonal, and thus are identical to each other. Numerically, for translation from parent to child, the translated expansion is then filtered via the procedure described in Section 10.6, so as to reduce the number of grid points at which it is discretized to that used at the child's level, discarding those parts of the far field signature which have insignificant effects on the child box.

10.15 Far Field to Local Translation

The following theorem provides a means for translating a far field expansion of the high-frequency diagonal form to a local expansion in the same form.

Theorem 20. Suppose $f_1 : S^2 \rightarrow \mathbb{C}$ is a far field signature centered at a point c_1 and valid outside a ball D_1 of center c_1 and radius R_1 ; denote the potential it yields by $\Phi : \mathbb{R}^3 \setminus D_1 \rightarrow \mathbb{C}$. Suppose D_3 is another ball, which has center c_3 and radius R_3 , such that $D_3 \cap D_1 = \emptyset$. Let the functions $\mu_k : S^2 \rightarrow \mathbb{C}$, for positive integer k , be defined by the formula

$$\mu_k(s) = \sum_{m=0}^k i^m \cdot (2m+1) \cdot P_m(\cos(\theta(c_3 - c_1, s))) \cdot h_m(\omega \|c_3 - c_1\|), \quad (10.48)$$

where $\theta(c_3 - c_1, s)$ denotes the angle between the vector $c_3 - c_1$ and the direction s . Let $g_k : S^2 \rightarrow \mathbb{C}$ be given by the formula

$$g_k(s) = f_1(s) \cdot \mu_k(s), \quad (10.49)$$

and let the functions $g_{k,n} : S^2 \rightarrow \mathbb{C}$ be defined by the formula

$$g_k(s) = \sum_{n=0}^{\infty} g_{k,n}(s), \quad (10.50)$$

together with the restriction that $g_{k,n}(s)$ must be a spherical harmonic of degree n . Let $\Phi_k : D_3 \rightarrow \mathbb{C}$ be given by the formula

$$\Phi_k(r, s) = \sum_{n=0}^{\infty} g_{k,n}(s) j_n(\omega r), \quad (10.51)$$

where r and s are coordinates in the system of spherical coordinates whose origin is c_3 . Then, for any $x \in D_3$,

$$\lim_{k \rightarrow \infty} \Phi_k(x) = \Phi(x). \quad (10.52)$$

Furthermore,

$$\max_{D_3} |\Phi_k(x) - \Phi(x)| = O \left(\left(\frac{R_1 + R_3}{\|c_3 - c_1\|} \right)^k \cdot \|f_1\| \right). \quad (10.53)$$

The above theorem is proven in [27]. The proof is based on the same formula (the addition theorem for spherical Bessel functions) on which Theorem 17 is based; accordingly, the truncation error, as given by (10.53), behaves similarly to the truncation error of partial-wave expansions (described above in Section 10.9): the significance of the k 'th term in the series is indicated by the magnitude of the product

$$h_k(\omega r) \cdot j_k(\omega \rho), \quad (10.54)$$

with

$$r = \|c_3 - c_1\|, \quad (10.55)$$

$$\rho < R_1 + R_3, \quad (10.56)$$

In the context of the FMM, far-field-to-local translation is done between boxes on the same level of the hierarchy which are separated by at least one box of the same size; thus the bounds (10.55), (10.56) become

$$r \geq 4R, \quad (10.57)$$

$$\rho < 2\sqrt{3}R, \quad (10.58)$$

where R is half the length of each of the boxes. Comparison of (10.58) to the corresponding bound (10.35) for partial-wave expansions shows that for sufficiently high frequencies, the appropriate truncation points to achieve a given accuracy, in the context of the FMM, are related as follows: when

translating a far field signature of degree p between boxes on the same level, the appropriate truncation point in (10.48) is $2p$. At lower frequencies, it is somewhat less than that; in the wideband FMM, it is determined numerically by evaluation of the worst-case value of (10.54), which is obviously

$$h_k(4\omega R) \cdot j_k(2\sqrt{3}\omega R). \quad (10.59)$$

With the above truncation point, (10.49) is a product of two sums of spherical harmonics, one of maximum degree p , the other of maximum degree $2p$. The resulting product is thus a sum of spherical harmonics of maximum degree $3p$. That product is then to be filtered, eliminating all but components of degree p or less; this can be done if the grid on which the product is sampled is of the type defined in Section 10.5, with at least $2p$ points in θ and $4p$ points in ϕ . Thus, numerically, the process of translating a far field expansion to a local expansion consists of resampling the far field expansion on a twice-finer grid, multiplying by the function μ_{2p} at each grid point, then filtering the result back onto the original grid, which yields the local expansion.

Three remarks on the numerical implementation of diagonal forms of translation operators seem in order.

Remark 3. The evaluation of the Far-Field-to-Local translation operators is where numerical instability manifests itself if an attempt is made to use high-frequency expansions for boxes that are too small. Indeed, the series (10.48) obviously does not converge as $k \rightarrow \infty$; in fact, $h_m(\omega \|c_3 - c_1\|)$ starts growing in magnitude, once m exceeds $\omega \|c_3 - c_1\|$. For sufficiently large $\omega \|c_3 - c_1\|$ (or, equivalently, for sufficiently large R), this problem does not occur, since the series is truncated before $h_m(\omega \|c_3 - c_1\|)$ becomes too large. As a practical matter, given a certain precision of calculations and a certain desired accuracy, one can determine the minimum size of the box for which the diagonal form of the translation operator provided by Theorem 20. When the calculations are performed in 64-bit floating-point arithmetic, Table 10 lists some of the cut-off points. We refer the reader to [25] for an excellent discussion of this class of issues.

Remark 4. The filtering and interpolation operations on the sphere are quite expensive compared to multiplication by a diagonal translation matrix, and at first glance appear to be the dominant element of the algorithm (so far as the CPU time requirements are concerned). Fortunately, each of them only needs to be performed once for each box: in the entire wideband FMM, there is only one function being interpolated for each box (the far

Table 10: Transition points to diagonal forms for translation operators

Requested Accuracy	Box Size (wavelengths)
10^{-3}	0.25
10^{-6}	3.50
10^{-9}	12.0

field expansion), and it is interpolated to the same grid each time. Likewise, as regards filtering, a number of functions, in the above descriptions, get filtered then added together to produce the local expansion for each box; but, since filtering is a linear operation, they can be added together before filtering, and the filter applied once.

Remark 5. The final remark concerns the calculation of the function μ_{2p} on grid points on the sphere. This calculation would be overly expensive if done naively: direct evaluation of (10.48) at each of $O(p^2)$ points would require $O(p^3)$ time. At the highest levels in the hierarchy, p is, in many important cases, on the order of $\sqrt{N}$ (where N is the total number of input points, as in (10.1)); thus a single $O(p^3)$ operation would result in an $O(N^{3/2})$ algorithm. But the only place that s enters into the calculation of $\mu_{2p}(s)$ is in the taking of the angle between s and the vector $c_3 - c_1$; thus μ_{2p} can be tabulated as a function of that angle, then interpolated as necessary for points on the sphere; if a local interpolation method is used for this, the tabulation takes $O(p^2)$ time ($O(p)$ points, with $O(p)$ time per point), as does the interpolation (to $O(p^2)$ points, with $O(1)$ time per point).

10.16 Exponential Expansions

For translation from far field expansions to local expansions in the low-frequency regime (that is, from h-expansions to j-expansions), the wideband FMM uses the scheme introduced in [20]. In that scheme, the translation is performed by means of “exponential expansions” (also known as “plane wave expansions”), which are based on the following formula for the Green’s function of the Helmholtz equation, which is valid for $z > 0$:

$$\frac{e^{i\omega r}}{r} = \frac{1}{2\pi} \int_0^\infty e^{-\sqrt{\lambda^2 - \omega^2} \cdot z} \int_0^{2\pi} e^{i\lambda(x \cos \alpha + y \sin \alpha)} d\alpha \frac{\lambda}{\sqrt{\lambda^2 - \omega^2}} d\lambda. \quad (10.60)$$

In [20], the outer integral in (10.60) was divided into two parts, a “propagating” part and an “evanescent” part, on each of which a different quadrature

was used. In the wideband FMM, we instead use a single quadrature for the whole of the outer integral. To determine the appropriate quadrature, we first evaluate the inner integral analytically, which transforms the right hand side of (10.60) into

$$\frac{1}{2\pi} \int_0^\infty e^{-\sqrt{\lambda^2 - \omega^2} \cdot z} J_0(\lambda \sqrt{x^2 + y^2}) \frac{\lambda}{\sqrt{\lambda^2 - \omega^2}} d\lambda; \quad (10.61)$$

an appropriate quadrature is thus one which integrates (10.61) accurately for any x , y , and z within the ranges used by the algorithm. Those ranges are as follows:

$$L \leq z \leq 4L, \quad (10.62)$$

$$-4L \leq x, y \leq 4L, \quad (10.63)$$

where L is the length (on each side) of the box to which the exponential expansion belongs.

The appropriate quadrature obviously depends on two variables: the wavenumber ω and the size L of the box. However, it depends chiefly on the product ωL (the real part of which is proportional to the size of the box in wavelengths); it is easy to show that if this is held constant while ω and L are varied, then the number of quadrature nodes required to achieve the same relative accuracy remains the same, while the weights and nodes are rescaled by a common factor.

The wideband FMM uses quadratures custom-tailored to this exact problem using a variant of the algorithm presented in [10]. These quadratures have a constant term; that is, (10.60) is approximated via the formula

$$\frac{e^{i\omega r}}{r} \approx a + \frac{1}{2\pi} \sum_{k=1}^s w_k e^{-\sqrt{\lambda_k^2 - \omega^2} \cdot z} \int_0^{2\pi} e^{i\lambda_k (x \cos \alpha + y \sin \alpha)} d\alpha \frac{\lambda_k}{\sqrt{\lambda_k^2 - \omega^2}}, \quad (10.64)$$

in which the accuracy of approximation is selected to be on the same order as the accuracy desired of the wideband FMM as a whole. The quadrature-generation algorithm takes as input that desired accuracy, and yields the number s of weights w_k and nodes λ_k , as well as the weights and nodes themselves and the constant a , to satisfy (10.64) to that accuracy. These quadratures are not computed at runtime, but are precomputed for each of several distinct ranges of the product ωL , which together encompass all the values of that product which are possible in the low-frequency regime.

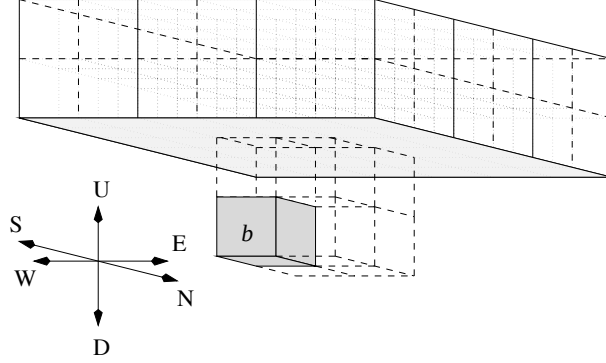
The inner integral of (10.60) is obviously an integral of a smooth, periodic function; thus the trapezoidal rule is appropriate for it. The amount of oscillation of the function to be integrated obviously depends on the value of the coefficient λ_k , as does thus the number of nodes required for the trapezoidal rule to achieve the desired accuracy; we denote that number by M_k , and denote the total number of quadrature nodes needed by S_{exp} ; clearly, $S_{exp} = 1 + \sum_{k=1}^s M_k$. A somewhat involved analysis, which we omit, shows that $s \approx p$, where p is the order of multipole expansion required to achieve the same accuracy, and also that $S_{exp} \approx p^2$.

The ranges (10.62) - (10.63) arise from the way exponential expansions are used, which is as follows. In the FMM, the “interaction list” of a box is defined to be the list of boxes on the same level of the hierarchy which are not adjacent to it (adjacency being defined as having any common boundary point, even just a corner), but whose parents are adjacent to its parent. The number of such boxes is clearly at most $6^3 - 3^3 = 189$ — fewer for boxes near the edge of the problem domain, or near larger boxes which were not subdivided to their level. For each box F in the interaction list of a box B , the FMM applies a far-field-to-local translation operator to convert the far field expansion on B into a local expansion on F .

To use exponential expansions for those translations, the interaction list of a box is partitioned into six lists, one for each face of the box:

- the $+z$ -list: boxes separated by at least one box in the $+z$ direction,
- the $-z$ -list: boxes separated by at least one box in the $-z$ direction,
- the $+y$ -list: boxes separated by at least one box in the $+y$ direction, and not contained in the $+z$ or $-z$ lists,
- the $-y$ -list: boxes separated by at least one box in the $-y$ direction, and not contained in the $+z$ or $-z$ lists,
- the $+x$ -list: boxes separated by at least one box in the $+x$ direction, and not contained in the $+z$, $-z$, $+y$, or $-y$ lists,
- the $-x$ -list: boxes separated by at least one box in the $-x$ direction, and not contained in the $+z$, $-z$, $+y$, or $-y$ lists.

A $+z$ -list is depicted in Figure 41. An exponential expansion as described above is referred to as facing in the $+z$ direction (that being the direction in which the exponentials in it decay). Exponential expansions facing in the other five directions are defined similarly, the only difference being in

Figure 41: The $+z$ -list of the box b

the names of the coordinate axes. For each of the six lists, an exponential expansion facing in that direction is used for translations to boxes in that list. It is evident that if a point $(x_0, y_0, z_0) \in \mathbb{R}^3$ is in a box B of length L , and a point $(x, y, z) \in \mathbb{R}^3$ is in a box of B 's $+z$ -list, then $L \leq z - z_0 \leq 4L$, $-4L \leq x - x_0 \leq 4L$, and $-4L \leq y - y_0 \leq 4L$; thus those are the ranges for which quadratures are constructed, as described above.

Substituting those quadratures into (10.60), the potential at a point (x, y, z) due to a source at a point (x_0, y_0, z_0) can be approximated via the formula

$$\begin{aligned} & \frac{e^{i\omega \|(x,y,z)-(x_0,y_0,z_0)\|}}{\|(x,y,z)-(x_0,y_0,z_0)\|} \\ & \approx a + \sum_{k=1}^s w_k e^{-\sqrt{\lambda_k^2 - \omega^2} \cdot (z - z_0)} \cdot \frac{\lambda_k}{\sqrt{\lambda_k^2 - \omega^2}} \cdot \\ & \quad \cdot \frac{1}{M_k} \sum_{j=1}^{M_k} e^{i\lambda_k((x-x_0) \cos \alpha_{kj} + (y-y_0) \sin \alpha_{kj})}, \end{aligned} \quad (10.65)$$

provided that $L \leq z - z_0 \leq 4L$, $-4L \leq x - x_0 \leq 4L$, and $-4L \leq y - y_0 \leq 4L$. Each of the terms summed up on the right hand side of (10.65) is evidently a plane wave; that is to say, it is of the form

$$W_j e^{(C_j, u - u_0)}, \quad (10.66)$$

where $u = (x, y, z)$ and $u_0 = (x_0, y_0, z_0)$, with $W_j \in \mathbb{R}$ and $C_j \in \mathbb{C}^3$. Using this notation, if sources at points $u_1, \dots, u_n$, with respective source strengths

$s_1, \dots, s_n$, are located in a box B , then the resulting potential at any point u located in any of the boxes of B 's $+z$ -list can be written as

$$\Phi(u) = \sum_{k=1}^n s_k e^{i\omega \|u-u_k\|} \approx \sum_{k=1}^n s_k \sum_{j=1}^J W_j e^{(C_j, u-u_k)}. \quad (10.67)$$

This, in turn, can be broken up into three parts: first, the creation of of an exponential expansion located at a point v_1 , via the formula

$$\alpha_j = W_j \sum_{k=1}^n s_k e^{(C_j, v_1-u_k)}, \quad (10.68)$$

for each $j = 1, \dots, J$, then (optionally) its translation to another point v_2 , via the formula

$$\beta_j = \alpha_j e^{(C_j, v_2-v_1)}, \quad (10.69)$$

for each $j = 1, \dots, J$, then its evaluation at the point u :

$$\Phi(u) \approx \sum_{j=1}^J \beta_j e^{(C_j, u-v_2)}. \quad (10.70)$$

The translation (10.69) can obviously be performed any number of times, with no loss of accuracy; the accuracy depends only on the relative locations of each pair of source and measurement points, which must satisfy the conditions (10.62), (10.63). In the wideband FMM, the location v_1 of the exponential expansion for a box B is chosen to be the center of the box, and the location v_2 from which the exponential expansion is evaluated is chosen to be the center of the box of B 's $+z$ -list that u resides in. With these choices, the creation operator (10.68) and the evaluation operator (10.70) are transposes of each other (thus possessing symmetry of the type described in Section 10.3), except that the former operator contains the weights $\{W_j\}$; having each operator contain the square roots of those weights would render the symmetry complete.

The advantage of using exponential expansions is, as previously mentioned, that their translation operator (10.69) is diagonal, and thus can be applied in $O(p^2)$ time. Since translation is the most frequently applied operation, being applied up to 189 times per box, this yields significant savings. The ability of exponential expansions to be translated more than once, with no loss of accuracy nor change of region of validity, allows for further savings:

examining the $+z$ -lists of the eight children of a parent box, and referring to the $+z$ direction as being “upwards”, it is evident that the $+z$ -lists of the top four children are the same, as are the $+z$ -lists of the bottom four children; and the former is a subset of the latter. Thus the $+z$ translations of those eight children can be accomplished by translating the exponential expansions of the bottom four children to a common point, adding them together, evaluating the resulting expansion on the boxes which are only in the bottom four children’s $+z$ -list, then translating the exponential expansions of the top four children to the same common point and adding them in, then evaluating the resulting expansion on the boxes which are in all eight children’s $+z$ -lists. This optimization reduces the maximum number of translations per box from 189 to roughly 40.

10.17 Conversion Between Exponential and Partial-Wave Expansions

In the wideband FMM, exponential expansions for each box are created not via (10.68) but rather from the h-expansion for that box; likewise, exponential expansions are not evaluated on a box via (10.70), but rather are converted into a j-expansion on that box. Since exponential expansions and partial-wave expansions each possess the symmetry described in Section 10.3, these two types of conversions (h-expansion to exponential expansion and exponential expansion to j-expansion) are transposes of each other; thus we describe only the latter type of conversion.

The representation of a plane wave as a partial-wave expansion belongs to a well-studied part of classical mathematical physics; the formula used in the wideband FMM for the conversion is easily derived from the following formula from [24] (vol. 2, 11.3.46):

$$e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{n=0}^{\infty} (2n+1) i^n \sum_{m=0}^n \epsilon_m \frac{(n-m)!}{(n+m)!} \cos(m(\phi - \nu)) \cdot P_n^m(\cos u) P_n^m(\cos \theta) j_n(kr), \quad (10.71)$$

where the vector $\mathbf{r}$ is of length r and has the spherical angles θ, ϕ , where the vector $\mathbf{k}$ is of length k and has the spherical angles u, v , and where ϵ_m is the Neumann factor ($\epsilon_m = 1$ when $m = 0$, $\epsilon_m = 2$ when $m > 0$). Using

the formula $\cos z = (e^{iz} + e^{-iz})/2$, (10.71) becomes

$$e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{n=0}^{\infty} (2n+1) i^n \sum_{m=-n}^n \frac{(n-|m|)!}{(n+|m|)!} e^{im(\phi-\nu)} \cdot P_n^{|m|}(\cos u) P_n^{|m|}(\cos \theta) j_n(kr). \quad (10.72)$$

To convert (10.72) to the form used in this paper, we make the substitutions $k = \omega$, $\nu = \alpha_{kj}$, and $\cos u = (i\sqrt{\lambda_k^2 - \omega^2})/\omega$, the last of which yields the formula $\sin u = \lambda_k/\omega$. The Cartesian components of the vector $\mathbf{k}$ which result from these substitutions are:

$$k_x = k \sin u \cos \nu = \lambda_k \cos \alpha_{kj}, \quad (10.73)$$

$$k_y = k \sin u \sin \nu = \lambda_k \sin \alpha_{kj}, \quad (10.74)$$

$$k_z = k \cos u = i\sqrt{\lambda_k^2 - \omega^2}. \quad (10.75)$$

Making the above substitutions into (10.76) yields the formula

$$e^{-\sqrt{\lambda_k^2 - \omega^2} \cdot z} \cdot e^{i\lambda_k(x \cos \alpha_{kj} + y \sin \alpha_{kj})} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \beta_n^m Y_n^m(\theta, \phi) j_n(\omega r), \quad (10.76)$$

in which (x, y, z) are the Cartesian equivalents of the spherical coordinates (r, θ, ϕ) , and where

$$\beta_n^m = 4\pi i^n \bar{P}_n^{|m|} \left(\frac{i\sqrt{\lambda_k^2 - \omega^2}}{\omega} \right) e^{im\alpha_{kj}}. \quad (10.77)$$

The left hand side of (10.76) is a single term of an exponential expansion (facing in the $+z$ direction), corresponding to the quadrature nodes λ_k and α_{kj} ; the right hand side is the equivalent j-expansion, with (10.77) being the formula for the coefficients of that expansion.

Denoting the coefficients of the exponential expansion by $\{E_{kj}\}$, the formula for converting the entire exponential expansion into a j-expansion is thus

$$\beta_n^m = \sum_{k=1}^s 4\pi i^n \bar{P}_n^{|m|} \left(\frac{i\sqrt{\lambda_k^2 - \omega^2}}{\omega} \right) \sum_{j=1}^{M_k} E_{kj} e^{im\alpha_{kj}}. \quad (10.78)$$

Since the nodes $\alpha_{k1}, \dots, \alpha_{kM_k}$ are equispaced, the inner sums of (10.78) can be performed using the FFT; doing so reduces the CPU time required for the whole conversion to $O(p^3)$.

The exponential expansion to which the above conversion formula applies faces in the $+z$ direction; thus, using the formula (10.78) on an exponential expansion facing in another direction results in a j -expansion relative to a system of coordinates whose z axis points in that direction; the latter must then be rotated (as described in Section 10.10) to the standard orientation.

10.18 Algorithm

The wideband FMM, as implemented, is adaptive; that is, the hierarchical subdivision of the problem domain is deeper in places where there are more source points: the rule used is that a box is subdivided if it encloses more than a certain number d of source points. (The number d is chosen so as to roughly minimize the CPU time.) However, in the high-frequency regime, no adaptivity is used: any box above the high-frequency cutoff line which contains any points whatsoever is subdivided. This removes the necessity to consider the cases of nearby boxes of different size which interact, one or more of those boxes being above the high-frequency cutoff line; instead, all such cases are entirely in the low-frequency regime. (Since the normal practice is to discretize objects using more than two points per wavelength, this is not a major limitation of the code.) After the problem domain is subdivided, the main computation is performed. It consists of three stages: first, creation of far field expansions for each box; second, translation of far field expansions to local expansions; third, evaluation of the local expansions.

The first stage is a bottom-up pass through the hierarchy of boxes. For each childless box, an h -expansion is created via (10.32). Then, for each parent box below the high-frequency cutoff line, an h -expansion is created from the h -expansions of its children, using the "point-and-shoot" method described in Section 10.11. At the high-frequency cutoff line, each h -expansion is converted into a far field signature via a spherical harmonic transform (Section 10.6). Above the cutoff line, far field signatures for parents are created from the far field signatures of their children as described in Section 10.14. At the end of this stage, a far field expansion (either an h -expansion or a far field signature) has been computed for all the boxes in the hierarchy.

In the second stage, for each box below the high-frequency cutoff line, six exponential expansions are created from the h -expansion for that box, as described in Section 10.17. These are then translated to each box in its interaction list, as described in Section 10.16. The resulting six exponential expansions (of the local-expansion variety) on each box are converted into j -expansions on that box, as described in Section 10.17. This accounts for all interactions between boxes of the same size in the low-frequency regime.

Interactions between boxes of different sizes only occur when the larger of those two boxes has not been subdivided, and thus has $O(1)$ points on it; accordingly, such interactions are handled by evaluating the smaller box's h-expansion directly on the larger box, or (for interactions going the other direction) by creating a j-expansion on the smaller box directly from the source points on the larger box. (In all cases, interactions are handled at as high a level as possible; thus interactions between boxes of different size only occur when the parent of the smaller box is closer than its own length to the larger box, and thus its expansions are invalid on the larger box.) For boxes above the high-frequency cutoff line, the far field signature for each box is translated to each of the boxes in its interaction list via the procedure described in Section 10.15.

The third stage is a top-down pass through the hierarchy of boxes. Starting from the top level, the local expansion on each parent box is translated into a local expansion on each of its children, and added to the existing local expansion (from the second stage) on that child. Above the high-frequency cutoff line, the translation process is performed as described in Section 10.14; to cross the line, a spherical harmonic transform is used; below the line, the translation process is performed as described in Section 10.11. After this, the j-expansion on each childless box is evaluated to yield the potential at the points in that box.

The above three stages handle all parts of (10.1) except for the interactions between points $\{x_j\}$ which are in childless boxes which are adjacent to each other, and between points $\{x_j\}$ which are in the same childless box; those interactions are evaluated directly.

10.19 Numerical Results

The wideband FMM has been applied to several test cases. The first of these is an aircraft-shaped object (Figure 42), 50 wavelengths in size. The surface of the aircraft was divided into 706,300 triangles, on each of which a single node was placed; the size of the smallest triangle was 1.06×10^{-6} wavelengths, and the size of the largest was 2.86×10^{-1} wavelengths. The wideband FMM was run at each of three different levels of accuracy, and the error calculated by comparing the results to those produced by direct application of (10.1). Table 11 shows the results, as well as the time (in seconds) taken by the direct method; the "error" columns contain the relative error in the L^2 sense (the L^2 norm of the error, divided by the L^2 norm of the correct result), and the first column contains the time taken by the direct evaluation of (10.1). For comparison, the results of applying

the FMM for the Laplace potential to the same geometry are shown in Table 12; the Laplace FMM used was the one described in [8], which is at a level of technology similar to that of the wideband FMM of this paper: it is adaptive, and it uses exponential expansions for diagonal translations.

Table 11: Example 1: Aircraft-shaped object

Time (direct)	Requested Accuracy	Error in potential	Error in gradient	Time (seconds)	Memory (MB)
337329	10^{-3}	0.43E-3	0.56E-3	485	300
337329	10^{-6}	0.48E-6	0.50E-6	1291	790
337329	10^{-9}	0.11E-9	0.95E-10	2947	1143

Table 12: Aircraft-shaped object - Laplace potential

Time (direct)	Requested Accuracy	Error in potential	Error in gradient	Time (seconds)	Memory (MB)
60590	10^{-3}	0.27E-3	0.37E-4	48.3	211
60590	10^{-6}	0.19E-6	0.43E-7	119	292
60590	10^{-9}	0.85E-10	0.61E-11	2437	376

Another example to which the wideband FMM was applied is a horse (Figure 43), also 50 wavelengths in size. The surface of the horse was divided into 872,694 triangles, on each of which a single node was placed; the size of the smallest triangle was 9.34×10^{-3} wavelengths, and the size of the largest one was 3.27×10^{-1} wavelengths. The results are depicted in Table 13, whose columns have the same meanings as the corresponding columns in Table 11. Again, the FMM for the Laplace potential was applied to the same geometry; the results are shown in Table 14.

The FMM of this paper was also applied to points on the surface of a sphere 50 wavelengths in size. The surface of the sphere was divided into 619,520 triangles (the smallest being 4.91×10^{-2} wavelengths in size, and the largest 6.27×10^{-2} wavelengths), on each of which a single node was placed. The results are tabulated in Table 15; the results for the Laplace potential in the same geometry are shown in Table 16.

Table 13: Example 2: Horse

Time (direct)	Requested Accuracy	Error (potential)	Error (gradient)	Time (seconds)	Memory (MB)
646143	10^{-3}	0.65E-3	0.31E-3	672	549
646143	10^{-6}	0.66E-6	0.92E-7	1832	1111
646143	10^{-9}	0.33E-9	0.33E-11	3515	2027

Table 14: Horse - Laplace potential

Time (direct)	Requested Accuracy	Error (potential)	Error (gradient)	Time (seconds)	Memory (MB)
107833	10^{-3}	0.91E-3	0.57E-3	63.7	328
107833	10^{-6}	0.46E-6	0.31E-6	139.7	322
107833	10^{-9}	0.25E-9	0.10E-9	298	584

Finally, we applied the FMM to points on the surface of a cube 50 wavelengths in size. The surface of the cube was divided into 619,520 triangles (each being 9.12×10^{-2} wavelengths in size), on each of which a single node was placed. The results can be found in Table 17; the results for the Laplace potential in the same geometry are shown in Table 18.

The orders of the various expansions used for boxes of various sizes, and for three different levels of requested accuracy ε , are listed in the Tables 19, 20. Table 19 lists the orders of partial-wave expansion used, in the columns labeled “LF”, and the orders of truncation of far-field-to-local translation operators (the number k in (10.48)), in the columns labeled “HF”. In each case the number of terms in the expansion is on the order of the square of the order: a partial-wave expansion of order p has $(p+1)^2$ terms; and, for translation, a far field signature is sampled on a grid of size roughly $2p \times 4p$ (the latter number being generally increased slightly so as to enable FFTs of that size to be performed efficiently), while the tabulated orders of truncation are, at high frequencies, roughly equal to $2p$. When a number is in parentheses (which only occurs in the “LF” column), it indicates that partial-wave expansions of that order are not used in the wideband FMM; instead, the high-frequency version is used.

The numbers of terms used in exponential expansions, for boxes of vari-

Table 15: Sphere

Time (direct)	Requested Accuracy	Error (potential)	Error (gradient)	Time (seconds)	Memory (MB)
324381	10^{-3}	0.27E-3	0.19E-3	521	416
324381	10^{-6}	0.15E-6	0.42E-7	1358	914
324381	10^{-9}	0.91E-10	0.24E-10	2873	1474

Table 16: Sphere-Laplace potential

Time (direct)	Requested Accuracy	Error (potential)	Error (gradient)	Time (seconds)	Memory (MB)
52936	10^{-3}	0.79E-3	0.90E-3	45	245
52936	10^{-6}	0.33E-6	0.45E-6	97.7	244
52936	10^{-9}	0.19E-9	0.12E-9	223	402

ous sizes, are listed in Table 20. (Each number is slightly larger than strictly necessary, since each of the numbers M_k which are summed to yield each entry in the table is adjusted upwards slightly to be a product of small primes, so that the FFT can be performed efficiently.)

The following observations can be made from the results of the numerical experiments described above, and from the more extensive experimentation we have performed.

1. The observed CPU times are compatible with the $n \cdot \log(n)$ estimate. One apparent exception is the CPU time for the 9-digit calculation for the aircraft-shaped object in Table 12. In fact, the FMM algorithm for the Laplace equation (we used the version from [8]) ran out of physical memory, and we see the CPU time penalty associated with the use of the virtual memory (the so-called paging). Using a computer with more memory or a more carefully written memory allocation for the Laplace FMM would eliminate the anomaly.
2. For larger boxes, the number of terms in the partial wave expansions (and, consequently, the cost of applying the translation operators) is almost independent of the requested precision (see Table 19); this is a well-known aspect of the behavior of such expansions. The number of terms in the exponential

Table 17: Cube

Time (direct)	Requested Accuracy	Error (potential)	Error (gradient)	Time (seconds)	Memory (MB)
376950	10^{-3}	0.97E-3	0.74E-3	393	364
376950	10^{-6}	0.73E-6	0.26E-7	1022	1295
376950	10^{-9}	0.23E-9	0.17E-10	2077	1001

Table 18: Cube - Laplace potential

Time (direct)	Requested Accuracy	Error (potential)	Error (gradient)	Time (seconds)	Memory (MB)
56433	10^{-3}	0.94-3	0.60E-3	52	201
56433	10^{-6}	0.41E-6	0.34E-6	132	272
56433	10^{-9}	0.28E-9	0.17E-9	231	362

expansions is also remarkably insensitive to the accuracy requirements (see Table 20).

3. In terms of CPU time requirements of the algorithm, it is advantageous to switch to the high-frequency (diagonal) form as soon (for cubes as small) as the accuracy considerations permit. When calculations are conducted in double precision (64 bit) arithmetic, and the answer is desired with three digits, the transition can be made very early (for boxes only 1/4 of a wavelength in size. When 9-digit accuracy is required, the transition is pushed to boxes that are as large as 12 wavelengths. This aspect of the algorithm is the principal reason for the relatively high cost of the scheme when higher accuracy is required.

4. The scheme tends to be about 10 times slower than the scheme for the Laplace equation in the same geometry. A factor of about 3 can be accounted for by the fact that the Laplace code uses predominantly real arithmetic, while the Helmholtz code is mostly complex. The remaining difference is related to the structure of the Helmholtz algorithm, principally at the boundary between the low and the high-frequency regimes.

5. As expected, both the time requirements and the accuracy of the algorithm are fairly insensitive to the nature of the charge distribution. Virtually

Table 19: Orders of partial-wave expansions

box size (wavelengths)	Expansion Order					
	$\varepsilon = 10^{-3}$		$\varepsilon = 10^{-6}$		$\varepsilon = 10^{-9}$	
	LF	HF	LF	HF	LF	HF
≤ 0.1	10		22		28	
0.25	11	7	23		38	
1	(15)	17	26		41	
2	(22)	32	31		46	
3.5	(31)	50	39	56	52	
5	(40)	68	(48)	77	59	
10	(70)	126	(80)	137	91	
12.5	(85)	155	(95)	166	107	176
15		183		195		206
20		240		253		265
30		352		367		380
40		463		480		494
50		575		593		608

arbitrary accuracies can be obtained (within the limitations of one's computational environment).

6. The scheme of this paper permits scattering problems involving hundreds of thousands of unknowns on the boundaries of the scatterers to be handled on modern desk-top computers, though the resulting CPU (and more importantly, wall-clock) times can be inconveniently long. Using more powerful (but still very accessible) systems, one can easily solve problems involving many millions of unknowns.

The reader might have observed that in the numerical examples in this paper, the size of scattering objects is limited by about fifty wavelengths. The reason for this limitation is that at about this point, the desktop computer used by the authors (Pentium 4 with 1.5 gigabytes of memory, using double precision arithmetic) tends to run out of memory. With other computer systems, considerably larger-scale scattering problems have been solved (see, for example, [30]).

Table 20: Numbers of terms in exponential expansions

box size (wavelengths)	Number of Exponential Terms		
	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-9}$
$< 10^{-4}$	415	1251	2576
10^{-4}	415	1251	2576
10^{-3}	415	1251	2575
10^{-2}	427	1271	2570
10^{-1}	426	1337	2561
0.5	722	1742	3149
1	898	2134	3646
2	1991	3407	5261
3	3540	5261	7156
4	5487	7720	9767
5	7619	10522	12662
6	10665	12785	16191
7	13860	16363	19499
8	18511	20399	24433
9	23445	24963	28774
10	28032	30071	33815

10.20 Conclusions

We have presented a wideband version of the Fast Multipole Method for the Helmholtz equation in three dimensions. Although the method has considerable internal complexity, it does not expose that complexity to the user by breaking down in any regime. It has asymptotic CPU time $O(N \log N)$, and, as demonstrated by numerical examples, even at high accuracies delivers very substantial speed increases over the direct method — more than two orders of magnitude — at problem sizes which fit on an ordinary personal computer at the time of this writing.

References

- [1] M. Abramowitz and I. Stegun (editors), “Handbook of Mathematical Functions”, National Bureau of Standards, 1964

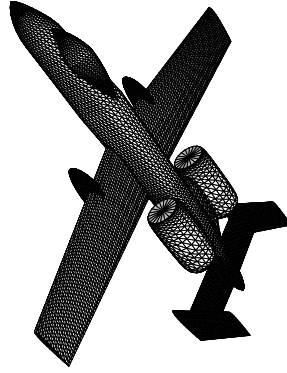


Figure 42: Surface triangulation of an aircraft-shaped object 50 wavelengths in size. The size of the smallest triangle is 1.06×10^{-6} wavelengths, and the size of the largest is 2.86×10^{-1} wavelengths.

- [2] B. Alpert, G. Beylkin, R. Coifman, V. Rokhlin, "Wavelet-like Bases for the Fast Solution of Second Kind Integral Equations," SIAM Journal of Scientific and Statistical Computing, V. 14, No. 1, January 1993, pp. 159-184.
- [3] G. Beylkin, R.R. Coifman, and V. Rokhlin, "Fast Wavelet Transforms and Numerical Algorithms I", Comm. Pure and Appl. Math., XLIV, 1991, 141-183.
- [4] L. C. Biedenharn and J. D. Louck, "Angular Momentum in Quantum Physics: Theory and Application", Addison-Wesley, Reading, 1981
- [5] E. Bleszinsky, M. Bleszinsky, T. Jaroszewicz, "AIM: Adaptive Integral Method for Solving Large-Scale Electromagnetic Scattering and Radiation Problems", Radio Science, v. 31, pp. 1225-1251, 1996.
- [6] N.N. Bojarski, "K-space Formulation of the Electromagnetic Scattering Problems", Air Force Avionic Lab. Technical Report AFAL-TR-71-75, 1971.



Figure 43: Surface triangulation of a horse 50 wavelengths in size. The size of the smallest triangle is 9.34×10^{-3} wavelengths, and the size of the largest is 3.27×10^{-1} wavelengths.

- [7] W. C. Chew, J. M. Jin, E. Michielssen, and J. Song (editors), “Fast and Efficient Algorithms in Computational Electromagnetics”, Artech House, 2001
- [8] H. Cheng, L. Greengard, and V. Rokhlin, “A Fast Adaptive Multipole Algorithm in Three Dimensions”, *Journal of Computational Physics* 155, 468-498 (1999).
- [9] R. Coifman, V. Rokhlin, and S. Wandzura, “The Fast Multipole Method for the Wave Equation: A Pedestrian Prescription”, *IEEE Antennas and Propagation Magazine*, Vol. 35, No. 3, 1993.
- [10] H. Cheng, V. Rokhlin, and N. Yarvin, “Non-linear Optimization, Quadrature, and Interpolation”, *SIAM Journal on Optimization*, Vol. 9, No. 4, pp. 901-923, 1999
- [11] R. Courant and D. Hilbert, “Methods of Mathematical Physics”, Interscience Publishers, 1953

- [12] E. Darve, "The Fast Multipole Method I: Error Analysis and Asymptotic Complexity", *SIAM Journal of Numerical Analysis*, v. 38, No. 1, pp. 98-128, 2000.
- [13] E. Darve, P. Have, "The Fast Multipole Method Numerical Implementation", *Journal of Computational Physics* (2000) 160, 195-240.
- [14] E. Darve, "A Fast Multipole Method for Maxwell's Equations Stable at All Frequencies", *Phil. Trans. R. Soc. Lond. A* (2004) 362, 1-27
- [15] B. Dembart, M. Epton, B.H.L. Fong, L. Greengard, S. Jiang, J.J. Ottusch, V. Rokhlin, J.L. Visser, S.M. Wandzura, "Advanced Electromagnetic Modeling: Final Technical Report", HRL REF K0006, Computational Physics Department, HRL Laboratories, December 2003.
- [16] M.A. Epton, B. Dembart, "Multipole Translation Theory for Three-Dimensional Laplace and Helmholtz Equations", *SIAM J. Sci. Comput.*, Vol. 16, No. 4, pp. 865-897, 1995.
- [17] L. Greengard and J. Huang, "A New Version of the Fast Multipole Method for Screened Coulomb Interactions in Three Dimensions", *Journal of Computational Physics*, v.180, pp 642-658 (2002)
- [18] L. Greengard and V. Rokhlin, "A Fast Algorithm For Particle Simulations", *J. Comput. Phys.*, vol. 73, pp. 325-348, 1987
- [19] L. Greengard, V. Rokhlin, "A New Version of the Fast Multipole Method for the Laplace Equation in Three Dimensions", *Acta Numerica*, 1997, pp. 229-269.
- [20] L. Greengard, J. Huang, V. Rokhlin, and S. Wandzura, "Accelerating Fast Multipole Methods for the Helmholtz Equation at Low Frequencies," *IEEE Computational Science and Engineering*, 5: (3), pp. 32-38, Jul-Sept 1998.
- [21] T. Hrycak, V. Rokhlin, "An Improved Fast Multipole Algorithm for Potential Fields", *SIAM Journal of Scientific Computing*, Vol. 19, No. 6, pp. 1804-1826, 1998.
- [22] R. Jakob-Chien and B. Alpert, "A Fast Spherical Filter with Uniform Resolution", *Journal of Computational Physics*, Vol. 136, No. 2, September 15, 1997, p 580-584

- [23] L.J. Jiang and W.C. Chew, "A Mixed-Form Fast Multipole Algorithm", Research Report: CCEM No. 03-05, Center for Computational Electromagnetics and Electromagnetics Laboratory, University of Illinois, 2005; to appear in *IEEE Transactions on Antennas and Propagation*.
- [24] P. M. Morse and H. Feshbach, "Methods of Theoretical Physics", McGraw-Hill, New York, 1953.
- [25] N. Nishimura, "Fast Multipole Accelerated Boundary Integral Equation Methods", Applied Mechanics Reviews, Vol. 55, No. 4, July 2002.
- [26] V. Rokhlin, "Rapid solution of integral equations of scattering theory in two dimensions," Journal of Computational Physics, vol. 86, pp. 414-439, 1990.
- [27] V. Rokhlin, "Diagonal forms of translation operators for the Helmholtz equation in three dimensions," Applied and Computational Harmonic Analysis, vol. 1, pp. 82-93, Academic Press, San Diego, 1993.
- [28] J.M. Song, W.C Chew, "Multilevel Fast Multipole Algorithm for Solving Combined Field Integral Equations of Electromagnetic Scattering", Microwave and Optical Technology Letters, Vol. 10, No. 1, September, 1995.
- [29] J. Stoer and R. Bulirsch, "Introduction to Numerical Analysis", Second Edition, Springer-Verlag, 1993.
- [30] S. Velamparambil, W. C. Chew, and J. M. Song, "10 million unknowns, is it that large", IEEE Antennas and Propagation Magazine, vol.45, no.2, pp.43-58, April 2003.
- [31] N. Yarvin and V. Rokhlin, "An Improved Fast Multipole Algorithm for Potential Fields on One-Dimensional Structures", SIAM Journal on Numerical Analysis, Vol. 36, No. 2, pp.629-666 (1999)
- [32] L. Ying, G. Biros, D. Zorin, "A Kernel-Independent Adaptive Fast Multipole Algorithm in Two and Three Dimensions", Journal of Computational Physics, vol. 196, pp. 591-626, 2004.
- [33] E. Yip, B. Dembart, 13-th Annual Review of Progress in Applied Computational Electromagnetics at the Naval Postgraduate School, Monterey, Ca, March 17, 1997.