A Super-Algebraically Convergent, Windowing-Based Approach to the Evaluation of Scattering from Periodic Rough Surfaces

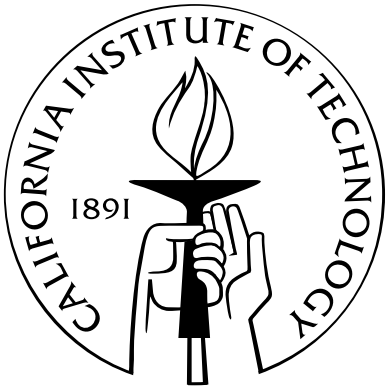
Thesis by

John A. Monro, Jr.

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2007

(Defended October 3, 2007)

To my grandparents and parents,
who taught me faith and a love of learning,
and to Lesley,
whose worth is far above rubies.

Acknowledgements

There have been so many friends and colleagues who have helped me in the process of creating this thesis and obtaining my Ph.D. degree. First of all, I thank my adviser, Professor Oscar Bruno. He gave to me his time, his intelligence and his enthusiasm, so that I could savor the highs of research, grow in my understanding of both mathematics and mathematical writing and produce satisfying results. His guidance and patience throughout the past few years have been vital to my success in this endeavor. I also wish to thank the other faculty and the staff of Caltech with whom I have had the privilege of interacting, especially the Applied and Computational Mathematics department's very kind and helpful administrator, Sheila Shull, and the rest of my thesis committee, Professors Dan Meiron, Houman Owhadi and Niles Pierce.

I am additionally grateful for the financial support provided via the Herbert Kunzel Fellowship during my first two years of graduate school as well as the support via Caltech teaching assistantships and grants from the Air Force Office of Scientific Research (AFOSR) and the National Science Foundation (NSF) during my remaining years.

There have been a number of friends who have greatly encouraged me and have enriched my life while I have been a graduate student, particularly during times of struggle. I thank them all, and I wish to name some of them here: Amos Anderson, David and Laura Barksdale, Nawaf Bou-Rabee, Ted and Katie Brenner, Shannon Browne, Roger Donaldson, Mark Furnish, Emilio Graff, Scott and Lai Holder, John Holdsworth, Walter and Azin Jones, Tracy Kidd, Wendy Mercer, Will Prentice, Eric Shrader, Marian Smith-Subbarao, Gary Velis and Andy Westhead.

I am grateful for my grandfather, a math teacher who taught me from a very young age to love math, and for my mother, who not only took care of me when I was young but also was the one who encouraged me to apply to Caltech as an undergraduate many years ago. Neither of them are with me today, but I think that they would be thrilled to see the

fruition of this work. I also am thankful for my grandmother, who helped inspire me to enjoy the English language and who has continued to be a blessing in my life. And I am glad that my father always has been there for me.

Finally, I wish to acknowledge and praise my wife, Lesley. She truly is an excellent wife: a support, a comfort and a friend.

Abstract

We introduce a new second-kind integral equation method to solve direct rough surface scattering problems in two dimensions. This approach is based, in part, upon the bounded obstacle scattering method that was originally presented in [12] and is discussed in an appendix of this thesis. We restrict our attention to problems in which time-harmonic acoustic or electromagnetic plane waves scatter from rough surfaces that are perfectly reflecting, periodic and at least twice continuously differentiable; both sound-soft and sound-hard type acoustic scattering cases—correspondingly, transverse-electric and transverse-magnetic electromagnetic scattering cases—are treated. Key elements of our algorithm include the use of infinitely continuously differentiable windowing functions that comprise partitions of unity, analytical representations of the integral equation's solution (taking into account either the absence or presence of multiple scattering) and spectral quadrature formulas. Together, they provide an efficient alternative to the use of the periodic Green's function found in the kernel of most solvers' integral operators, and they strongly mitigate the rapidly increasing computational complexity that is typically borne as the frequency of the incident field increases.

After providing a complete description of our solver and illustrating its usefulness through some preliminary examples, we rigorously prove its convergence. In particular, the super-algebraic convergence of the method is established for problems with infinitely continuously differentiable scattering surfaces. We additionally show that accuracies within prescribed tolerances are achieved with fixed computational cost as the frequency increases without bound for cases in which no multiple reflections occur.

We present extensive numerical data demonstrating the convergence, accuracy and efficiency of our computational approach for a wide range of scattering configurations (sinusoidal, multi-scale and simulated ocean surfaces are considered). These results include favorable comparisons with other leading integral equation methods as well as the non-

convergent Kirchhoff approximation. They also contain analyses of sets of cases in which the major physical parameters associated with these problems (i.e., surface height, wavenumber and incidence angle) are systematically varied. As a result of these tests, we conclude that the proposed algorithm is highly competitive and robust: it significantly outperforms other leading numerical methods in many cases of scientific and practical relevance, and it facilitates rapid analyses of a wide variety of scattering configurations.

Contents

Acknowledgements iv

Abstract vi

List of Figures xiii

List of Tables xviii

1 Introduction 1

1.1 Background . . . 2

1.1.1 Approximation-Based Methods . . . 3

1.1.2 Differential Equation Methods . . . 5

1.1.3 Integral Equation Methods . . . 5

1.2 Overview . . . 8

2 A New Algorithm for Periodic Rough-Surface Scattering Problems 10

2.1 Preliminaries . . . 11

2.1.1 Helmholtz and Maxwell Equations . . . 11

2.1.2 Rayleigh Expansion . . . 14

2.1.3 Integral Equations for the Scattering Cases . . . 15

2.1.3.1 Periodic Green's Function . . . 15

2.1.3.2 Integral Equations . . . 17

2.1.4 Computation of Scattering Efficiencies . . . 20

2.2 Periodic Green's Function vs. Partition of Unity . . . 22

2.2.1 Forms of the Periodic Green's Function . . . 23

2.2.2 Recasting the Integral Equations and Applying a Partition of Unity 25

2.2.3 Preliminary Simplified Case Study . . . 27

2.3 Two Representations of the Density . . . 35
2.3.1 First Representation of the Density . . . 35
2.3.2 A Second Pair of Equations and an Alternate Unknown . . . 36
2.3.2.1 High-Frequency Asymptotic Series Ansatz and Its Physical Restrictions . . . 37
2.3.2.2 Second Representation of the Density . . . 39
2.3.3 Physical Considerations in the Choice of Representations . . . 40
2.3.3.1 Types of Scattering . . . 41
2.3.3.2 Test for Multiple Reflections . . . 43
2.3.3.3 Test for Shadowing . . . 46
2.3.3.4 Height-to-Period Ratio vs. Incidence Angle . . . 46
2.3.3.5 Examples Illustrating the Behavior of $\mu_1(x)$ and $\mu_2(x)$ . . . 52
2.3.3.6 Choice of Representations . . . 60
2.4 Combining Windowing and the Two Representations . . . 60
2.5 Numerical Method . . . 61
2.5.1 Spectrally Accurate Quadratures on Analytic Closed Curves . . . 62
2.5.2 Numerical Handling of the Integral Operators . . . 63
2.5.2.1 Partitioning of the Integral Operators . . . 64
2.5.2.2 Quadrature Rule . . . 67
2.5.3 Linear Systems of Equations . . . 69
2.5.3.1 Computing Quadrature Weights . . . 72
2.5.3.2 FFT Interpolation and GMRES . . . 74
2.5.3.3 $\mathcal{O}(1)$ Computations for Simple-Reflection Cases . . . 75
2.5.4 Computation of Scattering Efficiencies . . . 76

3 Proofs **77**

3.1 Properties of Analytical Approximation . . . 78
3.1.1 Convergence of Integral Operators and Solutions . . . 79
3.1.1.1 Convergence of Integral Operators . . . 80
3.1.1.2 Existence and Convergence of Solutions . . . 87
3.1.2 Regularity of Solutions for Smooth Gratings . . . 88

3.1.3 Super-Algebraic Convergence of the Integral Operators and Solutions for Smooth Gratings . . . 94
3.1.3.1 Preliminary Considerations . . . 97
3.1.3.2 Foundational Lemmas . . . 100
3.1.3.3 Proof of Theorem 3.1.9 . . . 110
3.1.3.4 Super-Algebraic Convergence of Solutions for Smooth Gratings . . . 126
3.2 Properties of Numerical Method . . . 126
3.2.1 Convergence in Mesh Size . . . 128
3.2.1.1 Convergence of Quadratures . . . 129
3.2.1.2 Existence and Convergence of Numerical Solutions . . . 131
3.2.1.3 Super-Algebraic Convergence for Smooth Gratings . . . 132
3.2.2 $\mathcal{O}(1)$ Computational Times for Simple-Reflection Cases . . . 134
3.2.2.1 $\mathcal{O}(1)$ Approximation Errors . . . 135
3.2.2.2 $\mathcal{O}(1)$ Numbers of Target and Quadrature Points . . . 139

4 Numerical Results **143**
4.1 Convergence . . . 144
4.1.1 Multiple-Reflection Cases . . . 145
4.1.1.1 No Shadowing . . . 145
4.1.1.2 Shadowing . . . 147
4.1.2 Simple-Reflection Case . . . 149
4.2 Comparisons with Other Numerical Methods . . . 152
4.2.1 Overview of Numerical Methods Used for Comparison . . . 152
4.2.1.1 Methods of [4] . . . 152
4.2.1.2 Method of [13] . . . 153
4.2.1.3 Method Using the Kirchhoff Approximation . . . 154
4.2.2 Deterministic Grating Surfaces . . . 154
4.2.2.1 Cases from [4] . . . 155
4.2.2.2 Multi-Scale Cases . . . 165
4.2.2.3 Other Cases . . . 173
4.2.3 Simulated 1-D Ocean Surfaces . . . 177

4.3 Dependence of Algorithm Parameters upon Physical Parameters 186
4.3.1 Dependence upon k 186
4.3.1.1 Simple-Reflection Cases 187
4.3.1.2 Multiple-Reflection Cases 190
4.3.2 Dependence upon h 192
4.3.3 Dependence upon θ 195
4.3.4 Wood Anomaly Sensitivity 195
4.3.4.1 Computations at Wood Anomaly Values 198
4.3.4.2 Computations as k Approaches Wood Anomaly Values . . 200

5 Conclusions and Future Work **206**
5.1 Conclusions 206
5.2 Future Work 207

A A treatment of bounded obstacles **211**
A.1 Introduction 211
A.2 Boundary Integral Formulation 212
A.2.1 Ansatz 213
A.2.2 High-Frequency Integral Equation 216
A.3 Localized Integration 216
A.3.1 Stationary Phase 217
A.3.2 Convergent High-Frequency Integrator 218
A.3.3 Shadow Boundaries 222
A.4 Spectral Implementation 222
A.4.1 High-Order Interpolation 222
A.4.2 Trapezoidal-Rule Integration 224
A.5 Numerical Results 225
A.6 Extension to Non-Convex Scatterers 227
A.7 Conclusions 229
A.8 Evaluation of Stationary Points 229

B Code Verification **231**

C Additional TM/Sound-Hard Results **236**

D Rounding Errors **238**

Bibliography **240**

List of Figures

2.1 Partition of unity $P_1(x, x', c, A) + P_2(x, x', c, A) = 1$ 27

2.2 Partition of unity function $P_1(0, x', 0.1, 10)$ for $x' \in [0, 10]$ 32

2.3 Case with only simple reflections (top), case with simple and multiple reflections (bottom left) and case with simple reflections, multiple reflections and shadowing (bottom right) 42

2.4 $x_c > x$ (left) and $x_c = x$ (right) shadowing cases 45

2.5 Instances of multiple reflections for $x_c > x$ 45

2.6 Plot of $\frac{\partial \phi_2(x_1,x_2)}{\partial x_2}$ with $\theta = -\frac{\pi}{6}$, $h = 0.5$ and $x_1 = 0.25$ 48

2.7 Plot of $\frac{\partial \phi_2(x_1,x_2)}{\partial x_2}$ with $\theta = -\frac{\pi}{6}$, $h \approx 0.179$ and $x_1 = 0.25$ 48

2.8 Plot of a ray and the grating profile with $\theta = -\frac{\pi}{6}$, $h \approx 0.179$, $x_1 = 0.25$ and $x_2 \approx 0.712$ 49

2.9 Plots of rays and the grating profile with $\theta < 0$ (left) and $\theta > 0$ (right) 49

2.10 Plots of the multiple reflection threshold $h_{mult}(\theta)$ (solid line) and the shadowing threshold $h_{shad}(\theta)$ (dashed line) for the grating profile $f(x) = \frac{h}{2} \cos\left(\frac{2\pi x}{L}\right)$ 51

2.11 Case 1: real parts, imaginary parts and Fourier amplitudes of $\mu_1(x)$ and $\mu_2(x)$ 54

2.12 Case 2: real parts, imaginary parts and Fourier amplitudes of $\mu_1(x)$ and $\mu_2(x)$ 55

2.13 Case 3: real parts, imaginary parts and Fourier amplitudes of $\mu_1(x)$ and $\mu_2(x)$ 56

2.14 Case 4: real parts, imaginary parts and Fourier amplitudes of $\mu_1(x)$ and $\mu_2(x)$ 57

2.15 Case 5: real parts, imaginary parts and Fourier amplitudes of $\mu_1(x)$ and $\mu_2(x)$ 58

2.16 Case 6: real parts, imaginary parts and Fourier amplitudes of $\mu_1(x)$ and $\mu_2(x)$ 59

2.17 $P_{\text{sp}}(x, x', c_{\text{sp}}, A_{\text{sp}})$ and $P_{\text{rp}}(x, x', c_{\text{sp}}, A_{\text{sp}}, c, A)$, with $c_{\text{sp}} = 0.01$, $A_{\text{sp}} = \frac{7}{8}L$, $c = 0.5$ and $A = 3L$ 65

2.18 Real and imaginary parts of the kernel $g(x,x')\frac{h(ku(x,x'))}{e^{iku(x,x')}}e^{ik\phi_2(x,x')}$ ($x = 0.15$) and density $\mu_2(x)$ for Case 2 of Section 2.3.3.5. These functions are from the TE/sound-soft integral equation of (2.100). 70

4.1 Efficiencies of the Regime 1 case with no shadowing 146

4.2 Efficiencies of the Regime 1 case with shadowing 147

4.3 Efficiencies of the Regime 2 case 150

4.4 Errors of the $n_t = 16,\ n_i = 16\times 2,\ A = 30$ solution (left), the $n_t = 16,\ n_i = 16\times 3,\ A = 20$ solution (middle) and the $n_t = 16,\ n_i = 16\times 3,\ A = 30$ solution (right) for the Regime 2 case 151

4.5 Fourier amplitudes for the $n_t = 16,\ n_i = 16\times 3,\ A = 30$ solution (left) and the $n_t = 48,\ n_i = 48\times 1,\ A = 30$ solution (right) for the Regime 2 case . . . 151

4.6 Efficiencies (top), errors of this work's "mod. acc." solutions ($\mu_2(x)$ on the middle-left and $\mu_1(x)$ on the middle-right), errors of the "mod. acc." solution produced by the method of [13] (bottom-left) and errors of the KA-based method's solution (bottom-right) for Example 1 161

4.7 Efficiencies (top), errors of this work's "mod. acc." solution (middle-left), errors of the "4 digit" solution produced by the method of [13] (middle-right) and errors of the KA-based method's solution (bottom) for Example 2 164

4.8 The multi-scale surface $y = f(x) = \frac{0.025}{2}\left[\cos(2\pi x) + 0.04\sin(50\pi x)\right]$ with incidence angle vector $\boldsymbol{\alpha} = (\sin(30^\circ), -\cos(30^\circ))$ 169

4.9 Plot of $\frac{\partial\phi_2(x,x')}{\partial x'}$ for $\theta = 30^\circ$ with $x = 0.5$ (the function is discontinuous at $x' = x$) 169

4.10 Efficiencies (top), relative errors of this work's "mod. acc." solution (second level), relative errors of the "mod. acc." solution produced by the method of [13] (third level) and relative errors of the KA-based method's solution (bottom) for the $\frac{k}{2\pi} = 10.5,\ \theta = 30^\circ$ multi-scale case 170

4.11 Efficiencies (top), relative errors of this work's "mod. acc." solution (second level), relative errors of the "mod. acc." solution produced by the method of [13] (third level) and relative errors of the KA-based method's solution (bottom) for the $\frac{k}{2\pi} = 1000.5,\ \theta = 30^\circ$ multi-scale case 171

4.12 The multi-scale surface's first derivative 173

4.13 Efficiencies (top), relative errors of this work's "mod. acc." solution (second level), relative errors of the "mod. acc." solution produced by the method of [13] (third level) and relative errors of the KA-based method's solution (bottom) for the $\frac{k}{2\pi} = 10.5,\ \theta = 85°$ multi-scale case 174

4.14 Efficiencies (top), relative errors of this work's "mod. acc." solution (second level), relative errors of the "mod. acc." solution produced by the method of [13] (third level) and relative errors of the KA-based method's solution (bottom) for the $\frac{k}{2\pi} = 1000.5,\ \theta = 85°$ multi-scale case 175

4.15 The magnitudes of the randomly generated Fourier coefficients (top), the simulated 1-D ocean surface (middle), the first derivative of the surface (bottom left) and the second derivative of the surface (bottom right) 179

4.16 Efficiencies for the $\theta = 5°$ case (top), plus relative errors of our method's $n_{\mathrm{t}} = 1280$ solution (second level), our method's $n_{\mathrm{t}} = 2304$ solution (third level) and the KA-based approach's $n_{ka} = 3840$ solution (bottom) 181

4.17 Efficiency-by-efficiency differences between our method's $n_{\mathrm{t}} = 1280$ solution and the KA-based $n_{ka} = 1280$ solution for the $\theta = 5°$ case 182

4.18 Efficiencies for the $\theta = 80°$ case (top), plus relative errors of our method's $n_{\mathrm{t}} = 768$ solution (middle) and the KA-based approach's $n_{ka} = 3840$ solution (bottom) 184

4.19 Efficiencies for the $\theta = 85°$ case (top), plus relative errors of our method's $n_{\mathrm{t}} = 768$ solution (middle) and the KA-based approach's $n_{ka} = 3840$ solution (bottom) 185

4.20 Fourier amplitudes for $\frac{k}{2\pi} = 100$ (left) and $\frac{k}{2\pi} = 1000$ (right) for $\theta = 10°$ and $h = 0.2$ 189

4.21 Fourier amplitudes for the $\mu_2(x)$, $n_{\mathrm{t}} = 24$ representation (left), $\mu_2(x)$, $n_{\mathrm{t}} = 2000$ representation (middle) and the $\mu_1(x)$ representation (right) for $\frac{k}{2\pi} = 1000$, $\theta = 10°$ and $h = 0.2$ 190

4.22 Fourier amplitudes for the $\mu_1(x)$ representation (left) and the $\mu_2(x)$ representation (right) for $\frac{k}{2\pi} = 1000$, $\theta = 10°$ and $h = 0.25$ 191

4.23 Zoomed plot of the multiple reflection threshold (solid line) and the shadowing threshold (dashed line) as a function of θ for the grating profile $f(x) = \frac{h}{2}\cos\left(\frac{2\pi x}{L}\right)$ 196

4.24 Set 1: Fourier amplitudes of $\mu_1(x)$ (left) and $\mu_2(x)$ (right) for $\frac{k}{2\pi} = 9.5$ (top), $\frac{k}{2\pi} = 10$ (middle) and $\frac{k}{2\pi} = 10.5$ (bottom) 201

4.25 Set 2: Fourier amplitudes of $\mu_1(x)$ (left) and $\mu_2(x)$ (right) for $\frac{k}{2\pi} = 9.5$ (top), $\frac{k}{2\pi} = 10$ (middle) and $\frac{k}{2\pi} = 10.5$ (bottom) 202

4.26 Set 3: Fourier amplitudes of $\mu_1(x)$ (left) and $\mu_2(x)$ (right) for $\frac{k}{2\pi} = 99.5$ (top), $\frac{k}{2\pi} = 100$ (middle) and $\frac{k}{2\pi} = 100.5$ (bottom) 203

A.1 Scattering by a circular cylinder $\boldsymbol{r}(\theta) = (a\cos\theta, a\sin\theta)$, $0 \le \theta \le 2\pi$, with $ka = 20$; all graphs show real and imaginary parts of complex quantities as functions of the angular coordinate θ. Top left: $\varphi(\boldsymbol{r}(\theta))$ (solution of (A.6)). Top right: $\varphi(\boldsymbol{r}(\theta))/(ke^{ik\boldsymbol{\alpha}\cdot\boldsymbol{r}(\theta)})$. Bottom left: $\mu(\boldsymbol{r}(\theta)) = \frac{\partial\Psi(\boldsymbol{r}(\theta))}{\partial\boldsymbol{\nu}(\boldsymbol{r}(\theta))}$ (solution of (A.3)). Bottom right: $\mu(\boldsymbol{r}(\theta))/(ke^{ik\boldsymbol{\alpha}\cdot\boldsymbol{r}(\theta)})$. 215

A.2 Real part of functions $f_A(x)e^{ikx^p}$ and $f_\varepsilon(x)e^{ikx^p}$ with upper envelopes $f_A(x)$ and $f_\varepsilon(x)$, respectively; $p = 2$. 218

A.3 Circular scatterer under plane wave incidence: target point T ($\theta_0 = 0$) and stationary phase points S_1, S_2 and S_3 220

A.4 Circular scatterer under plane wave incidence: smooth cutoffs around the critical points for $\theta_0 = \pi/8$, with $k = 1000$ (top) and $k = 4000$ (bottom). The quantity displayed in both graphs is the real part of the integrand in (A.18), divided by $\cos(\theta)$. 221

A.5 Top left: shadow boundary for the circular cylinder. Top right: cubic root change of variables. Bottom left and bottom right: variation of the phase of $\mu_{\text{slow}}^{\text{exact}}$ at the shadowing point before and after application of the change of variables, respectively, for $k = 100$ to $k = 1000000$. 223

A.6 A multiple scattering configuration: the kite-shaped scatterer of equation (A.21) 227

A.7 Real parts of $\mu(\boldsymbol{r}(\theta))$ and $\mu(\boldsymbol{r}(\theta))/e^{ik\boldsymbol{\alpha}\cdot\boldsymbol{r}(\theta)}$ on the kite-shaped scatterer, for $k = 800$ 227

A.8 Real and imaginary parts of the slow densities $\mu_{\text{slow}}^0(\boldsymbol{r}(\theta))$ and $\mu_{\text{slow}}^1(\boldsymbol{r}(\theta))$ on the kite-shaped scatterer, for $k = 800$ 228

B.1 Case 3 efficiencies and the differences in their computed values between this work and the solver of [13] 232

B.2 Case 4 (top), Case 5 (middle) and Case 6 (bottom) efficiencies 234

D.1 Fourier amplitudes for the $A = 30$ solution (left) and the $A = 30000$ solution (right) for the simple-reflection case of Section 4.1.2 239

List of Tables

2.1 Approximation errors for various A ($\bar{k_n} = 2\pi$) 32

2.2 Approximation errors for various A ($\bar{k_n} = \frac{\pi}{2}$) 33

2.3 Physical quantities for the examples of this section 52

4.1 Convergence table for various A ($n_\mathrm{t} = n_\mathrm{i} = 192$) for the Regime 1 case with no shadowing 146

4.2 Convergence table for various n_t ($n_\mathrm{t} = n_\mathrm{i}$ and $A = 800$) for the Regime 1 case with no shadowing 146

4.3 Convergence table for various n_t and n_i ($A = 800$) for the Regime 1 case with no shadowing 147

4.4 Convergence table for various A ($n_\mathrm{t} = 96,\ n_\mathrm{i} = 96 \times 3$) for the Regime 1 case with shadowing 148

4.5 Convergence table for various n_t ($n_\mathrm{i} = n_\mathrm{t}$ and $A = 750$) for the Regime 1 case with shadowing 148

4.6 Convergence table for various n_t and n_i ($A = 750$) for Regime 1 case with shadowing 148

4.7 Convergence table for various A ($n_\mathrm{t} = n_\mathrm{i} = 48$) for the Regime 2 case 150

4.8 Convergence table for various A ($n_\mathrm{t} = 16$ and $n_\mathrm{i} = 16 \times 3$) for the Regime 2 case 150

4.9 Convergence table for various n_t ($n_\mathrm{i} = n_\mathrm{t}$ and $A = 30$) for the Regime 2 case 150

4.10 Convergence table for various n_i ($n_\mathrm{t} = 16$ and $A = 30$) for the Regime 2 case 151

4.11 Physical parameters for the cases that are described in both [4] and [21] . . . 157

4.12 Results of the four methods of [4] plus the Nyström method of [44] for Example 1 160

4.13 Results of this thesis' method for Example 1 160

4.14 Results of the method of [13] for Example 1 160

4.15 Result of the KA-based method for Example 1 161

4.16 Results of the four methods of [4] plus the Nyström method of [44] for Example 2 163
4.17 Results of this thesis' method for Example 2 163
4.18 Results of the method of [13] for Example 2 163
4.19 Result of the KA-based method for Example 2 164
4.20 Results of the four methods of [4] plus the Nyström method of [44] for Example 3 166
4.21 Results of this thesis' method for Example 3 166
4.22 Results of the method of [13] for Example 3 166
4.23 Result of the KA-based method for Example 3 166
4.24 Efficiencies and errors for Example 3. The errors listed for this work as well as for the method of [13] come from the two solvers' "mod. acc." solutions. . 166
4.25 Results for the multi-scale cases with $\theta = 30°$ using this work's method. The energy balance errors are listed as "e.b. error." 172
4.26 Results of the method of [13] for the multi-scale cases with $\theta = 30°$ 172
4.27 Result of the KA-based method for the multi-scale cases with $\theta = 30°$ 172
4.28 Results for the multi-scale cases with $\theta = 85°$ using this work's method ($\mu_1(x)$) 173
4.29 Results of the method of [13] for the multi-scale cases with $\theta = 85°$ 176
4.30 Result of the KA-based method for the multi-scale cases with $\theta = 85°$ 176
4.31 Further computational results 177
4.32 Table for the solutions for the $\theta = 5°$ case as computed by the method of this thesis 182
4.33 Table for the solutions for the $\theta = 5°$ case as computed by the KA-based approach 182
4.34 Table for the solutions for the $\theta = 80°$ case as computed by the method of this thesis. The time without computing the efficiencies was not determined for the $n_t = 1024$ reference case. 183
4.35 Table for the solution for the $\theta = 80°$ case as computed by the KA-based approach 183
4.36 Table for the solutions for the $\theta = 85°$ case as computed by the method of this thesis. The time without computing the efficiencies was not determined for the $n_t = 1024$ reference case. 183
4.37 Table for the solution for the $\theta = 85°$ case as computed by the KA-based approach 184

4.38 Table for $\theta = 10^\circ$ and $h = 0.025$ . . . 188
4.39 Table for $\theta = 10^\circ$ and $h = 0.2$ . . . 189
4.40 Table for $\frac{k}{2\pi} = 1000$, $\theta = 10^\circ$ and $h = 0.2$. The energy balance errors are listed as "e.b. error." . . . 189
4.41 Table for $\theta = 10^\circ$ and $h = 0.25$ . . . 191
4.42 Table for $\frac{k}{2\pi} = 1000$, $\theta = 10^\circ$ and $h = 0.25$ . . . 191
4.43 Results of the method of [13] ($\theta = 10^\circ$ and $h = 0.25$) . . . 191
4.44 Table for various h ($\frac{k}{2\pi} = 10$, $\theta = 10^\circ$) . . . 193
4.45 Table for various h using $\mu_1(x)$ for all cases ($\frac{k}{2\pi} = 10$, $\theta = 10^\circ$) . . . 194
4.46 Table for various h ($\frac{k}{2\pi} = 10$, $\theta = 60^\circ$) . . . 194
4.47 Table for various h using $\mu_1(x)$ for all cases ($\frac{k}{2\pi} = 10$, $\theta = 60^\circ$) . . . 194
4.48 Results of the method of [13] for the $h = 4.0$ cases ($\frac{k}{2\pi} = 10$) . . . 194
4.49 Table for various θ ($\frac{k}{2\pi} = 10$ and $h = 0.025$) . . . 196
4.50 Table for various θ using $\mu_1(x)$ for all cases ($\frac{k}{2\pi} = 10$ and $h = 0.025$) . . . 196
4.51 Table for various θ ($\frac{k}{2\pi} = 1000$ and $h = 0.025$) . . . 196
4.52 Table for various θ using $\mu_1(x)$ for all cases ($\frac{k}{2\pi} = 1000$ and $h = 0.025$) . . . 197
4.53 Results of the method of [13] ($h = 0.025$) . . . 197
4.54 The physical parameters for each Wood Anomaly case ($h = 0.025$) . . . 200
4.55 Table for the Set 1 cases, computed using $\mu_2(x)$ ($h = 0.025$, $\theta = 0^\circ$, $n_\mathrm{t} = 48$, $n_\mathrm{i} = 48 \times 1$) . . . 200
4.56 Table for the Set 2 cases, computed using $\mu_2(x)$ ($h = 0.025$, $\theta = 30^\circ$, $n_\mathrm{t} = 60$, $n_\mathrm{i} = 60 \times 1$) . . . 200
4.57 Table for the Set 3 cases, computed using $\mu_2(x)$ ($h = 0.025$, $\theta = 30^\circ$, $n_\mathrm{t} = 432$, $n_\mathrm{i} = 432 \times 1$) . . . 204
4.58 Table for various k ($h = 0.025$, $\theta = 30^\circ$, $n_\mathrm{t} = 60$, $n_\mathrm{i} = 60 \times 1$) . . . 204
4.59 Table for various k ($h = 0.025$, $\theta = 30^\circ$, $n_\mathrm{t} = 60$, $n_\mathrm{i} = 60 \times 1$) . . . 205

A.1 Localized integrator, sinusoidal slow density (error on $I(\theta_0 = 0)$ using N integration points) . . . 220
A.2 Interpolation of the slow density (number of coefficients in the Fourier expansion of $\mu_{\mathrm{slow}}^{\mathrm{exact}}$ for a circular scatterer, with and without the change of variables depicted in Figure A.5) . . . 222

A.3 Scattering of an incident plane wave on a circular cylinder of radius a 225

B.1 Physical quantities, numerical parameters and results for the cases computed by our method and the method of [13]. The energy balance errors are listed as "e.b. error." 231

B.2 Case 1 efficiencies and the differences in their computed values between this work and the solver of [13] 232

B.3 Case 2 efficiencies and the differences in their computed values between this work and the solver of [13] 232

B.4 Physical quantities, numerical parameters and results for the cases from [17] 233

B.5 Efficiencies and relative errors of Case 4 235

B.6 Efficiencies and relative errors of Case 5 235

B.7 Efficiencies and relative errors of Case 6 235

C.1 Physical quantities and choices of representation for the densities of the TM cases, plus the sections in Chapter 4 discussing the corresponding TE cases . 237

C.2 Numerical parameters for the TM cases 237

C.3 TE and TM energy balance errors 237

D.1 Results for various A ($n_{\mathrm{t}} = n_{\mathrm{i}} = 48$) for the simple-reflection case of Section 4.1.2238

Chapter 1

Introduction

Scattering theory describes the actions of particles and waves as they encounter objects and inhomogeneities in media. In classical physics this includes the propagation and scattering of light and sound waves by obstacles such as vehicles, biological bodies and buried objects, as well as the area of focus of this thesis: interactions of waves with rough surfaces. The diffraction of light by gratings in spectrometers, the reflection and absorption of music in concert halls, the propagation of radio waves over the ocean surface and the generation of radar clutter by land and water bodies are only just a small sampling of the diverse phenomena that comprise such interactions. Thus, rough surface scattering has been the subject of a rich literature, and it has a broad variety of applications in various fields including the military, medicine, communications, materials science, environmental science and many others.

Many problems in these fields are of the "direct" type: scattered waves are computed for given incident waves and scattering surfaces. Their solutions are of significant interest—forming, for example, a critical part of many engineering design cycles by allowing scattering effects to be modeled before prototypes are built and tested. In this thesis, we focus upon the evaluation of direct problems, developing a computational method with a view toward applicability in real-life contexts. Conversely, an "inverse" problem involves seeking a description of a scattering surface from a knowledge of scattered wave data generated by one or more incident waves. The discrimination of land mines from other buried objects is such a problem. We do not treat inverse problems in this thesis; we merely mention here that their solutions are usually computed by iterative processes involving direct problems of the types discussed in the present work.

Accuracy and efficiency are often of the highest priority in the solution of direct problems

because of stringent accuracy requirements and/or limitations in the acceptable running times for a given application. With these goals in mind, we introduce a numerical method that uses a second-kind integral equation to solve direct rough surface scattering problems in two dimensions. We specifically consider problems in which incident time-harmonic acoustic or electromagnetic plane waves impinge from above onto rough surfaces that are perfectly reflecting, periodic and at least twice continuously differentiable; both sound-soft and sound-hard type acoustic scattering cases—correspondingly, transverse-electric (TE) and transverse-magnetic (TM) electromagnetic scattering cases—are treated.

Two major difficulties in the construction of integral equation-based solvers are addressed by our approach, namely: 1) the evaluation of the periodic Green's function found in the kernel of most solvers' integral operators, and 2) the rapidly increasing computational complexity typically borne as the frequency (correspondingly, the wavenumber) of the incident field increases. We address these issues via consideration of certain smooth (i.e., C^∞) windowing functions that comprise partitions of unity (POU), certain analytical representations of the integral equation's solution and certain quadrature formulas. With regards to point 1), our method allows us to avoid the use of the periodic Green's function by evaluating the integral operator by means of a POU-based spectral quadrature rule we introduce that facilitates extremely efficient and accurate evaluations (this rule is super-algebraically convergent for problems with smooth scattering surfaces). As far as point 2) is concerned, on the other hand, the use of POUs allows us to localize the integration within small regions of the scattering surface, and in the absence of multiple scattering our method also incorporates a physically intuitive and particularly helpful representation of the solution—yielding accuracies within prescribed tolerances with computational cost that does not increase as the frequency (wavenumber) increases without bound. Thus, we obtain a highly competitive algorithm, convergent and robust, which significantly outperforms other leading numerical methods in many cases of scientific and practical relevance, as well as facilitates rapid analyses of a wide variety of scattering configurations.

1.1 Background

In 1907, Lord Rayleigh published his renowned paper [49] on electromagnetic scattering from diffraction gratings. After almost one hundred years of further investigation, much

progress has been made on the rough surface scattering problem. There is still, however, much fruitful research left to be done, e.g., in developing ever more accurate and efficient algorithms for scattering in two and three dimensions, in determining their domains of validity and in applying them to larger and more complicated physical situations such as high-frequency and low-grazing angle scattering.

Various computational approaches for rough surface scattering are surveyed in a recent paper [56]. In that overview, the authors classify all numerical methods as being "approximation-based methods," "differential equation methods" or "integral equation methods." We briefly review these categories here, highlighting just some of the methods described in that paper—along with other methods and techniques that are of particular note—so as to place our integral equation method into a broad research context.

1.1.1 Approximation-Based Methods

As noted in [56], all numerical methods involve making approximations. This particular category of methods contains those which utilize analytical approximations to the solutions of the Helmholtz and Maxwell equations rather than, say, finite differentiation, finite quadrature, etc.

A number of approximation-based methods are based upon asymptotic expansions in the frequency (wavenumber) of the incident field (taking into account the frequency relative to the scattering surface's roughness). Such expansions include the "small-perturbation method" (SPM), which is used for low frequencies. Also, there is the high-frequency "Kirchhoff approximation" (KA), which is based upon locally approximating the scattering surface by tangent planes. KA may or may not include a correction for shadowing (see Section 2.3.3.1 for a discussion about the various kinds of scattering, including multiple scattering and shadowing, that can occur in the problems we are investigating); with the correction, it is known as "physical optics" (PO), although sometimes the terms are used interchangeably without such a distinction being made. The infinite-frequency limit of PO is "geometrical optics" (GO), which treats the incident field as a set of rays which propagate in accordance with Fermat's principle (see, e.g., [7]). An important extension of GO is the "geometrical theory of diffraction" (GTD) by Keller (see, e.g., the classic paper [32]). GTD more properly handles shadow regions as well as singularities on the rough surface such as edges and corners (see Section 2.3.3.1 for further discussion regarding shadowing). SPM

and GO have been combined into "composite" or "two-scale" models to handle surfaces with multiple roughness scales; other extensions also have been made. The above analytical approximations have been incorporated into numerical algorithms, e.g., using KA to approximate a "density" function on the scattering surface (see Section 2.1.3 regarding density functions) and then integrating this quantity to determine the scattered field (see Section 2.3.2.1 for a discussion of this application of KA).

We make special mention of a related high-frequency method that is developed in [17]: an asymptotic series expansion of the density is developed using an ansatz in which the density is represented by a "rapidly oscillating" precomputed factor multiplying a presumably "slowly oscillating" periodic function that is determined using Taylor-Fourier series and asymptotic expansions of oscillatory integrals (see Section 2.3.2.1). It is similar to the GO-based high-frequency asymptotic series described in works such as [58], but, unlike that series, it does not involve the solution of a phase function via the eikonal equation. The ansatz from [17] has motivated one of the representations of the density used by our own algorithm, as we describe in Section 2.3.2.

Another important approximation-based method is based upon the Rayleigh expansion of the scattered field as a series of outgoing plane waves [49] (see Section 2.1.2 for a brief discussion of this expansion). While this expansion rigorously holds for all points above the top of the scattering surface (a fact we make use of in the computations we present in this thesis; see Section 2.1.4), the Rayleigh hypothesis assumes that the expansion also holds for all points on and above the surface. Using this assumption, which only holds under very restrictive conditions, a variety of methods have been devised to compute the coefficients of the expansion (see, e.g., [48]).

As noted in [56], these and other methods not listed here have proved themselves useful for a variety of applications. Yet, they are subject to the difficulties of determining the conditions under which they are theoretically valid and under which they are numerically useful (note that there is not necessarily a one-to-one correspondence in these properties!). For example, the KA-based method developed in [6] neglects multiple scattering and shadowing effects—which may arise as the incidence angle approaches grazing or as the slopes of the rough surfaces under consideration increase (see Section 2.3.3.1). Also, the region of validity of the high-frequency method in [17] does not include shadowing. Yet another example is the Rayleigh hypothesis: this approximation is sometimes accurate even where

the hypothesis theoretically fails (see, e.g., [42] for a discussion of the validity of the hypothesis). Furthermore, in addition to such validity concerns, there is the problem of certain methods becoming computationally intensive when high-order expansion terms are added to improve their accuracy. Nevertheless, approximation-based methods are still widely used today, and the physical insights gained from them continue to inspire the development of new "fully numerical" methods such as the one presented in this thesis.

1.1.2 Differential Equation Methods

Differential equation methods directly solve the differential equations of scattering in a volumetric mesh above the rough surface (and also below it, in cases for which the incident field is partially transmitted, e.g., electromagnetic scattering from dielectric materials). A variety of these algorithms exist, e.g., finite-difference time-domain methods (FDTD) and finite element methods (FEM). The matrices that arise in differential equation methods are sparse—thus, allowing for computational efficiencies in their generation and manipulation—since the differential operators are local, but the number of unknowns can be extremely large due to the volumetric meshes employed [56]. Algorithms such as FDTD and FEM also require the use of artificial boundary conditions away from the scattering surface in order to limit the size of the computational region, and this presents additional challenges. Thus, while differential equation methods are still being used for certain applications, they are not as commonly used as the approximation-based and integral equation methods.

1.1.3 Integral Equation Methods

Integral equation methods are the most commonly used numerical approaches to rough surface scattering [56]. For perfectly reflecting surfaces the differential equations of scattering (including their boundary conditions) are transformed into a single Fredholm-type integral equation, as is done in this thesis; a pair of coupled integral equations arise when transmission of the incident field occurs. The integral equations are solved for certain density functions defined on the scattering surface, which can then be used to yield the scattered field throughout space. Thus, integral equation methods are often preferred over differential equation methods, because they require surface discretizations instead of volume discretizations and because they automatically incorporate conditions of radiation at infinity. Also, integral equation methods are convergent for all scattering configurations (with a caveat

on certain wavenumbers associated with "Wood Anomalies" for periodic rough surfaces; we discuss this issue in various places throughout this thesis, beginning in Section 2.1.2), including configurations with multiple scattering and shadowing. Thus, they are significantly more reliable than approximation-based methods. See [19, 20, 35] for in-depth treatments of integral equations, their use in scattering theory and their numerical solutions—much of which is directly applicable to scattering from bounded obstacles but can also be applied to rough surface scattering.

As we stated earlier, for two-dimensional configurations involving perfectly reflecting surfaces there are two fundamental scattering cases, namely TE/sound-soft and TM/sound-hard (see Section 2.1.1), and multiple integral equation formulations exist for each of these types. These formulations involve one of two kinds of integral equations—commonly known as the "first-kind" and "second-kind" integral equations. First-kind equations are of the form

$$\int_{\partial D} K(x,x')\mu(x')\,ds(x') = g(x), \quad x \in \partial D, \tag{1.1}$$

where $K(x,x')$ is the kernel of the integral operator, $g(x)$ is a known function, $\mu(x)$ is the unknown density to be computed and ∂D is the surface of integration (in this context, the scattering surface). For scattering problems of the type we are considering, the first-kind integral equations that arise are ill-posed [35] and are often numerically treated through the use of preconditioning matrices which multiply the approximating linear system [56]. On the other hand, second-kind equations, such as the ones we use in this thesis, are of the form

$$\mu(x) - \int_{\partial D} K(x,x')\mu(x')\,ds(x') = g(x), \quad x \in \partial D. \tag{1.2}$$

Equations of this type do not suffer the type of numerical instabilities exhibited by first-kind equations, although preconditioners for increasingly rough surfaces have also been effectively used. First-kind and second-kind integral equations can be derived through the use of Green's second identity as well as by other means; see Section 2.1.3 for derivations of the particular integral equations (one per type of scattering case) that form the basis for our method.

Integral equation formulations of periodic rough surface scattering of the type we are examining typically contain what is known as the "periodic Green's function" in the kernels of their integral operators (see Section 2.1.3). This function can be represented as a series

with terms containing the two-dimensional free space Green's function for the Helmholtz equation, but many numerical methods, such as the methods of [4, 13, 21], use other forms of this function which are more computationally efficient. See Section 2.2.1 for a brief discussion of the forms of the periodic Green's function used in [4, 13, 21]; see [39] for a broader survey of methods for its evaluation. As stated previously, in this thesis we take a different approach in the formulation of the integral operator—one which avoids use of the periodic Green's function altogether—by using instead the free space Green's function and smooth POUs (see Section 2.2).

In deriving approximating linear systems for integral equations, a number of different types of discretization schemes have been implemented. One such scheme is the "Nyström method" for second-kind integral equations, which approximates the integral operator using a quadrature rule and computes an approximation to the density on a set of quadrature points. This approach to the rough surface scattering problem has been taken recently in the work described in [44], for example. Other schemes called "projection methods," e.g., the "collocation method" and the "Galerkin method," involve projecting the integral equation onto finite-dimensional subspaces. The collocation and Galerkin methods have been used for both first-kind and second-kind integral equations in the computational methods of [4, 13, 21]. An advantage of the Nyström approach is the computational ease with which the associated matrix elements can be evaluated; it is this fact which has lead us to adopt a Nyström approach in our work (see, e.g., [35] for further discussion of the Nyström and projection methods). Our numerical method is based upon the Nyström method for bounded obstacle scattering in two dimensions that is presented in [20], and for C^∞ scattering surfaces it yields super-algebraic convergence in the number of discretization points due to the spectrally accurate quadrature rule employed (see Sections 2.5 and 3.2). Additionally, we note that the Nyström methods of [20, 44] and the projection methods of [4, 21] approximate integral equations by linear systems with $N \times N$ matrices, thus computing $N \times N$ elements; our method, however, uses $N \times M$ elements (N is the number of points used to represent the density and M is the number of quadrature points) with N sometimes significantly less than M (the N-point density is Fourier interpolated to M points when evaluating the integral operator), which can result in significant computational savings.

Even though integral equation methods only discretize the scattering surface (as opposed to a volume like differential equation methods), the matrices which arise in their approx-

imating equations are dense (as opposed to the sparse matrices of differential equation methods), and treatments of these matrices can become computationally costly, particularly in the high-frequency regime. Iterative solvers of various types have been employed in order to handle this difficulty, including the "iterated Kirchhoff approximation" (i.e., computing the Neumann series solution for second-kind integral equations) and Krylov subspace methods, as well as "Fast" methods such as those introduced in [14, 52]. The Krylov subspace methods include the conjugate gradient method (CG), the quasi-minimum residual method (QMR) and the generalized minimum residual method (GMRES); GMRES [53] is often the most robust of the Krylov subspace methods, although it can become costly in terms of memory for slowly convergent problems unless re-starts are utilized. Many of the "Fast" methods incorporate the use of the "Fast Fourier Transform" (FFT), which is valuable for performing rapid interpolations and matrix-vector multiplications. Finally, for high-frequency scattering, the contributions to the density function at any point are dominated by nearby interactions on the scattering surface. The "banded matrix iterative algorithm" (BMIA) is one method which takes advantage of this insight. We conclude by mentioning the bounded obstacle scattering method described in [12], which uses certain representations for the density together with a quadrature rule involving smooth POU-based windowing functions that shrink with increasing wavenumber. This algorithm, which also uses GMRES and FFT-based interpolations for efficiency, has motivated, in part, the rough surface solver of this thesis—a solver that can be used for small or large wavenumbers. See Section 2.5 for a discussion of our quadrature method as well as our use of GMRES and FFTs; the work presented in [12] is included in this thesis as Appendix A.

1.2 Overview

As we have stated, in this thesis we introduce a new integral equation-based computational approach to the rough surface scattering problem. We begin by describing the method and illustrating its usefulness through some preliminary examples (Chapter 2). Chapter 3 contains proofs of theorems associated with various components of our method. Numerical results demonstrating the convergence, accuracy and efficiency of our computational approach for a variety of cases (sinusoidal, multi-scale and simulated ocean surfaces are considered) are given in Chapter 4. Finally, we state conclusions and some potential directions

of future research (Chapter 5), and we include an appendix discussing closely related work (which originally appeared in [12]) on scattering from bounded obstacles (Appendix A).

Chapter 2

A New Algorithm for Periodic Rough-Surface Scattering Problems

In this chapter we provide a complete description of our numerical algorithm for scattering by periodic rough surfaces. We begin with a review of the differential equations of scattering, and we establish certain corresponding integral equations and expressions for the scattered fields that are used throughout this thesis (Section 2.1). The role of the periodic Green's function associated with these integral equations is discussed, and in Section 2.2 an alternative approach, based upon the use of a smooth partition of unity (POU) in conjunction with our reformulated integral equations that does not require the use of periodic Green's functions, is developed. In this section, we also present a highly simplified example showing how the use of the POU-based approach could give rise to rapidly convergent approximations of the relevant integral operators. Two representations of the surface density are introduced in Section 2.3, and a discussion motivating their use under various scattering configurations is given: one of the formulations is shown to be preferable for cases in which no multiple scattering of the incident field occurs (especially for high-frequency problems), while the other is more appropriate when multiple scattering is present. In Section 2.4, our insights about the POU and the two density representations are combined to produce the particular formulations of the integral equations we solve, and, finally, in Section 2.5 we describe the Nyström-type approximation we employ which yields super-algebraic convergence in the number of discretization points given smooth scattering profiles.

The excellent convergence properties of the computational approximations that we introduce in this chapter are rigorously proved in Chapter 3. The numerical examples pre-

sented in Chapter 4, finally, demonstrate the qualities of our solver and present extensive comparisons with previous leading approaches.

2.1 Preliminaries

2.1.1 Helmholtz and Maxwell Equations

We consider the scattering of electromagnetic (EM) and acoustic waves by rough surfaces, that is, by interfaces between two homogeneous media in the three-dimensional rectangular coordinate system $Oxyz$. Our attention is restricted to electromagnetic scattering from perfectly conducting rough surfaces along with sound-soft and sound-hard acoustic scattering, with incident fields that propagate in directions parallel to the Oxy plane. Also, we only consider surface profiles that are described by equations of the form $y = f(x),\ x \in (-\infty, \infty)$ (independent of z), where, defining

$$C^r_{\text{per}}(L) \equiv \{f \in C^r : f \text{ is } L\text{-periodic}\}, \tag{2.1}$$

we consider $f \in C^r_{\text{per}}(L)$, i.e., $f(x)$ is r-times continuously differentiable and L-periodic, for some positive integer $r \geq 2$ or $r = \infty$. These conditions give rise to two-dimensional scattering systems; this is often referred to as scattering from one-dimensional rough surfaces.

Remark 2.1.1. *Scattering from surfaces with less differentiability, e.g., surfaces containing corners, can be treated by introducing appropriate changes of variables [20, 30], but such problems will not be pursued in this thesis.*

In particular, we seek to determine the scattered waves that result as time-harmonic plane waves traveling in the domain $y > f(x)$, which are denoted by

$$u^{\text{inc}}(\boldsymbol{r}, t) \equiv e^{i(k\boldsymbol{\alpha}\cdot\boldsymbol{r}-\omega t)} \equiv \psi^{\text{inc}}(\boldsymbol{r})e^{-i\omega t} \tag{2.2}$$

and

$$\boldsymbol{E}^{\text{inc}}(\boldsymbol{r}, t) \equiv \boldsymbol{A}e^{i(k\boldsymbol{\alpha}\cdot\boldsymbol{r}-\omega t)} \tag{2.3}$$

$$\boldsymbol{H}^{\text{inc}}(\boldsymbol{r}, t) \equiv \boldsymbol{B}e^{i(k\boldsymbol{\alpha}\cdot\boldsymbol{r}-\omega t)} \tag{2.4}$$

in the acoustic and electromagnetic cases, respectively, impinge upon the interface $y = f(x)$. Here, $k \equiv \frac{2\pi}{\lambda} > 0$ denotes the wavenumber (spatial frequency), λ is the wavelength, ω is the time-frequency,

$$\boldsymbol{\alpha} \equiv (\sin(\theta), -\cos(\theta)) \tag{2.5}$$

is the direction of propagation ($-\frac{\pi}{2} < \theta < \frac{\pi}{2}$, measured counterclockwise from the negative Oy-axis), $\boldsymbol{r} \equiv (x, y)$ and the three-dimensional vector constants $\boldsymbol{A} \equiv (A_x, A_y, A_z)$ and $\boldsymbol{B} \equiv (B_x, B_y, B_z)$, associated with the incident electric and magnetic fields, are related to each other and the direction of propagation according to the relations $\boldsymbol{A}\cdot(k\sin(\theta), -k\cos(\theta), 0) = 0$ and $\boldsymbol{B} = \frac{1}{\omega\mu_0}(k\sin(\theta), -k\cos(\theta), 0) \times \boldsymbol{A}$. Given our assumption on the properties of the material making up the scattering surface, these incident waves are perfectly reflected back into the domain $y > f(x)$. Thus, in the acoustic case, the scattered wave $\psi^{\text{scat}}(\boldsymbol{r})e^{-i\omega t}$, like the incident wave, obeys the two-dimensional Helmholtz equation [20]

$$\Delta\psi^{\text{scat}}(\boldsymbol{r}) + k^2\psi^{\text{scat}}(\boldsymbol{r}) = 0, \quad y > f(x) \tag{2.6}$$

(dropping the factor $e^{-i\omega t}$), satisfying either the Dirichlet boundary condition (for sound-soft scattering)

$$\psi^{\text{scat}}(\boldsymbol{r}) = -\psi^{\text{inc}}(\boldsymbol{r}), \quad y = f(x) \tag{2.7}$$

or the Neumann boundary condition (for sound-hard scattering)

$$\frac{\partial\psi^{\text{scat}}(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})} = -\frac{\partial\psi^{\text{inc}}(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})}, \quad y = f(x); \tag{2.8}$$

here,

$$\boldsymbol{\nu}(\boldsymbol{r}) \equiv \frac{(-f'(x), 1)}{|(-f'(x), 1)|} \tag{2.9}$$

is the upward normal to the surface at the point $\boldsymbol{r} = (x, f(x))$ ($|(x, y)| \equiv \sqrt{x^2 + y^2}$), and the normal derivative $\frac{\partial\psi^{\text{scat}}(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})}$ is taken as a limit from above the surface:

$$\frac{\partial\psi^{\text{scat}}(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})} \equiv \lim_{\epsilon\to+0} \nabla\psi^{\text{scat}}(\boldsymbol{r} + \epsilon\boldsymbol{\nu}(\boldsymbol{r}))\cdot\boldsymbol{\nu}(\boldsymbol{r}), \quad y = f(x). \tag{2.10}$$

Analogously, the incident and scattered waves obey the time-harmonic Maxwell equa-

tions [42]

$$\begin{aligned} \nabla \times \boldsymbol{E}^{\text{scat}} &= i\omega\mu_0 \boldsymbol{H}^{\text{scat}} \quad , \ \nabla \cdot \boldsymbol{E}^{\text{scat}} = 0 \\ \nabla \times \boldsymbol{H}^{\text{scat}} &= -i\omega\epsilon_0 \boldsymbol{E}^{\text{scat}} \quad , \ \nabla \cdot \boldsymbol{H}^{\text{scat}} = 0 \end{aligned} \quad , \quad y > f(x), \tag{2.11}$$

in the electromagnetic case ($k^2 = \omega^2\epsilon_0\mu_0$, $e^{-i\omega t}$ is again dropped and we assume that there are no free charge or free current densities in $y > f(x)$), where

$$\begin{aligned} \boldsymbol{\nu}(\boldsymbol{r}) \times \left(\boldsymbol{E}^{\text{scat}} + \boldsymbol{E}^{\text{inc}}\right) &= 0 \\ \boldsymbol{\nu}(\boldsymbol{r}) \cdot \left(\boldsymbol{H}^{\text{scat}} + \boldsymbol{H}^{\text{inc}}\right) &= 0 \end{aligned} \quad , \quad y = f(x). \tag{2.12}$$

It can be shown [42] that each component of $\boldsymbol{E}^{\text{scat}}$ and $\boldsymbol{H}^{\text{scat}}$ solves the Helmholtz equation (2.6) and that, in the present two-dimensional case, the full vector equations and boundary conditions can be reduced to two uncoupled scalar problems. One of these problems assumes a transverse electric (TE) polarization of the incident wave ($A_z = 1,\ A_x = A_y = B_z = 0$) and takes the transverse component E_z^{scat} of the scattered electric field as unknown; it is easy to check that E_z^{scat} satisfies the same problem as the sound-soft scattered field. Under transverse magnetic (TM) polarization, on the other hand, H_z^{scat} is the solution of the sound-hard problem if $B_z = 1,\ B_x = B_y = A_z = 0$.

To ensure a physically meaningful and unique solution $\psi^{\text{scat}}(\boldsymbol{r})$ of the TE/sound-soft and TM/sound-hard equations, a radiation condition is added [48]: the scattered field is bounded (in magnitude) and outward traveling as $y \to \infty$ (see Section 2.1.2 for a precise definition of this condition in the present context). Additionally, in view of the uniqueness of solutions, it can be shown (see, e.g., [42]) that $\psi^{\text{scat}}(\boldsymbol{r})$ is $\alpha \equiv k\sin(\theta)$ quasi-periodic—defining

$$(\alpha, -\beta) \equiv k\boldsymbol{\alpha} = (k\sin(\theta), -k\cos(\theta))\,, \tag{2.13}$$

we have $\psi^{\text{scat}}(x + L, y) = e^{i\alpha L}\psi^{\text{scat}}(x, y)$—just as $\psi^{\text{inc}}(\boldsymbol{r}) = e^{i\alpha x - i\beta y}$ is α quasi-periodic. Because of this property, it is sufficient to determine $\psi^{\text{scat}}(\boldsymbol{r})$ over a single L-length interval of x.

Remark 2.1.2. *Further discussion of the existence and uniqueness of the solutions can be found in [18]; in particular, this reference contains a proof of uniqueness for the TE/sound-soft problem.*

2.1.2 Rayleigh Expansion

The scattered field $\psi^{\text{scat}}(\boldsymbol{r})$ can be represented in $y > \max[f(x)]$ as a series of outgoing plane waves that is called the "Rayleigh Expansion." Following [42, 48], let

$$\alpha_n \equiv \alpha + n\frac{2\pi}{L} = k\sin(\theta) + n\frac{2\pi}{L}, \quad n \in \mathbb{Z}, \tag{2.14}$$

let U be the finite set of integers n such that $k^2 - \alpha_n^2 > 0$ and let

$$\beta_n \equiv \begin{cases} \sqrt{k^2 - \alpha_n^2} & , \quad n \in U \\ i\sqrt{\alpha_n^2 - k^2} & , \quad n \notin U, \end{cases} \tag{2.15}$$

where, in equation (2.15) as well as throughout this thesis, given a positive number a, the symbol $\sqrt{a}$ denotes the positive square root of a. We assume that for all n we have $k \neq \pm\alpha_n$, or, equivalently, $\beta_n \neq 0$; $k = \pm\alpha_n$ are wavenumbers corresponding to the physical phenomena known as "Wood Anomalies" [16, 50, 59]. Under this assumption, for $y > \max[f(x)]$, every α quasi-periodic solution to the Helmholtz equation (2.6) can be written as

$$\psi^{\text{scat}}(\boldsymbol{r}) = \sum_{n=-\infty}^{\infty} A_n e^{i\alpha_n x - i\beta_n y} + \sum_{n=-\infty}^{\infty} B_n e^{i\alpha_n x + i\beta_n y}. \tag{2.16}$$

In order to disallow incoming or unbounded waves as $y \to \infty$, we impose the following radiation condition:

Definition 2.1.1. *A solution in $y > \max[f(x)]$ of the form (2.16) is said to satisfy the radiation condition as $y \to \infty$ and is called "radiating" if $A_n = 0$ for all n. Similarly, such a solution in $y < \min[f(x)]$ satisfies the radiation condition as $y \to -\infty$ if $B_n = 0$ for all n.*

Thus, a radiating solution $\psi^{\text{scat}}(\boldsymbol{r})$ can be expanded in $y > \max[f(x)]$ as

$$\psi^{\text{scat}}(\boldsymbol{r}) = \sum_{n=-\infty}^{\infty} B_n e^{i\alpha_n x + i\beta_n y}. \tag{2.17}$$

Remark 2.1.3. *The waves corresponding to $n \in U$ are constant in magnitude and are called "propagating," whereas for $n \notin U$ the waves exponentially decay as $y \to \infty$ and are called "evanescent." The nth propagating wave travels in the direction $(\frac{\alpha_n}{k}, \frac{\beta_n}{k})$; the number*

of propagating waves is a locally constant function of k, which changes by either one or two units at the Wood Anomaly wavenumbers $k = \pm\alpha_n$.

2.1.3 Integral Equations for the Scattering Cases

As is well known, the Helmholtz and Maxwell problems discussed above can be recast in terms of integral equations containing what is known as the "periodic Green's function." For each of the TE/sound-soft and TM/sound-hard cases, there are a variety of integral equations that have been used—see, e.g., [55, Chapter 3] for an illustration of how multiple integral equation formulations are possible for the same case. The periodic Green's function, and the particular equations which will be the basis of all of our work in this thesis, are reviewed briefly here.

2.1.3.1 Periodic Green's Function

We begin by considering the radiating fundamental solution to the two-dimensional Helmholtz equation [20, 55]:

$$\Phi(\boldsymbol{r}, \boldsymbol{r}') \equiv \frac{i}{4} H_0^1(k|\boldsymbol{r} - \boldsymbol{r}'|), \tag{2.18}$$

where $\boldsymbol{r}' \equiv (x', y')$ and $H_0^1(z) = J_0(z) + iY_0(z)$ is the order zero Hankel function of the first kind ($J_0(z)$ and $Y_0(z)$ are the order zero Bessel functions of the first and second kind, respectively). The fundamental solution $\Phi(\boldsymbol{r}, \boldsymbol{r}')$, also known as the "free space Green's function," satisfies

$$\Delta\Phi(\boldsymbol{r}, \boldsymbol{r}') + k^2\Phi(\boldsymbol{r}, \boldsymbol{r}') = -\delta(\boldsymbol{r} - \boldsymbol{r}') \tag{2.19}$$

together with a radiation condition which imposes that the wave is bounded and outgoing as $|\boldsymbol{r}| \to \infty$ for compact sets of $\boldsymbol{r}'$. Thus [16, 42], the radiating α quasi-periodic fundamental solution of

$$\Delta\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}') + k^2\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}') = -\delta(y - y') \sum_{n=-\infty}^{\infty} e^{-i\alpha nL}\delta(x - x' + nL) \tag{2.20}$$

is

$$\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}') \equiv \frac{i}{4} \sum_{n=-\infty}^{\infty} e^{-i\alpha nL} H_0^1(k|(x - x' + nL, y - y')|), \tag{2.21}$$

or, equivalently for $(x-x', y-y') \neq (nL, 0)$ for any $n \in \mathbb{Z}$,

$$\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}') = \frac{i}{2L} \sum_{n=-\infty}^{\infty} \frac{e^{i\alpha_n(x-x')+i\beta_n|y-y'|}}{\beta_n}. \tag{2.22}$$

The function $\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')$ is the "periodic Green's function."

Remark 2.1.4. *Both $\Phi(\boldsymbol{r}, \boldsymbol{r}')$ and $\Phi_{per}(\boldsymbol{r}, \boldsymbol{r}')$ contain a logarithmic singularity at $\boldsymbol{r} = \boldsymbol{r}'$ due to the Bessel function of the second kind $Y_0(z)$ [3]. The treatment of this singularity is an important component of the algorithm presented in this thesis.*

Denoting

$$\begin{aligned} \mathcal{B} &= \mathcal{B}(x, y) \\ &\equiv \left\{ (x', y') \in \mathbb{C}^2 : -L - \gamma \leq \Re[x - x'] \leq L + \gamma, \ \left|\Im[x - x']\right| < \frac{L}{2}, \ \left|\Im[y - y']\right| < \frac{L}{2} \right\} \end{aligned} \tag{2.23}$$

for some small $\gamma > 0$ and letting

$$\mathcal{K} \equiv \{k \in \mathbb{C} : \Im[k] \geq 0, \ k \neq 0, \ k \neq \pm\alpha_n \text{ for every } n \in \mathbb{Z}\}, \tag{2.24}$$

each term of the series

$$\frac{i}{4} \sum_{n \neq -1,0,1} e^{-i\alpha n L} H_0^1(k|(x - x' + nL, y - y')|)$$

is well defined and differentiable for (x', y', k) in compact subsets of $\mathcal{B} \times \mathcal{K}$, and the series converges uniformly in such subsets [16] (related series which closely pertain to our algorithm are shown in Chapter 3 to converge uniformly). Thus, for an integral of the form

$$\int_{\mathcal{P}(x)} \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}') \mu(\boldsymbol{r}') \, ds(\boldsymbol{r}'),$$

where

$$\mathcal{P}(x) = \left\{ \boldsymbol{r}' = (x', y') : x - \frac{L}{2} < x' < x + \frac{L}{2}, \ y' = f(x') \right\} \tag{2.25}$$

and $\mu(\boldsymbol{r}')$ is some continuous function on $y' = f(x')$, $\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')$ is a "weakly singular" kernel with a log-type singularity [19]; the same is true of the normal derivatives of $\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')$.

We note that the above integral exists as an improper integral if $\boldsymbol{r} = (x, f(x))$.

Remark 2.1.5. *As stated in Section 2.1.1, in this thesis k is real and $k > 0$. In accordance with the discussion of Section 2.1.2, we say that k is a "Wood Anomaly value" if $k = \alpha_n$ or $k = -\alpha_n$ for some integer n, where α_n is given by (2.14); for completeness, this definition includes $k = 0$. The periodic Green's function $\Phi_{per}(\boldsymbol{r}, \boldsymbol{r}')$ is undefined for k which are Wood Anomaly values, since for each such k there is a value of n such that $\beta_n = 0$.*

2.1.3.2 Integral Equations

Let H_+ be a real number and let $H_+ > \max[f(x)]$,

$$D_+ = D_+(x, H_+) \equiv \left\{ \boldsymbol{r}' = (x', y') : x - \frac{L}{2} < x' < x + \frac{L}{2},\ f(x') < y' < H_+ \right\}, \tag{2.26}$$

$u(x', y') \equiv \psi^{\text{scat}}(\boldsymbol{r}')$, $v(x', y') \equiv \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')$ and $\boldsymbol{r} \in D_+$. To derive an integral representation for $\psi^{\text{scat}}(\boldsymbol{r})$ for the TE/sound-soft system, following, e.g., [42, Appendix B] we compute

$$\int_{D_+} \left\{ u \left[\Delta v + k^2 v \right] - v \left[\Delta u + k^2 u \right] \right\} d\boldsymbol{r}'$$

in two different ways. First, by the differential equations—the Helmholtz equation for u and (2.20) for v—the integral evaluates to $-\psi^{\text{scat}}(\boldsymbol{r})$. Second, by Green's second identity we have

$$\begin{aligned} \int_{D_+} \left[u \left(\Delta v + k^2 v \right) - v \left(\Delta u + k^2 u \right) \right] d\boldsymbol{r}' &= \int_{D_+} \left(u \Delta v - v \Delta u \right) d\boldsymbol{r}' \\ &= \int_{\partial D_+} \left(-u \frac{\partial v}{\partial \boldsymbol{\nu}(\boldsymbol{r}')} + v \frac{\partial u}{\partial \boldsymbol{\nu}(\boldsymbol{r}')} \right) ds(\boldsymbol{r}'), \end{aligned} \tag{2.27}$$

where $ds(\boldsymbol{r}')$ is the differential arc length and the usual sign conventions have been switched so that $\boldsymbol{\nu}(\boldsymbol{r}')$ is the internal normal to the boundary of D_+ (consistent with the earlier definition of $\frac{\partial \psi^{\text{scat}}(\boldsymbol{r})}{\partial \boldsymbol{\nu}(\boldsymbol{r})}$ on $y = f(x)$). Now, the integrals along

$$\left\{ \boldsymbol{r}' = (x', y') : x' = x - \frac{L}{2},\ f(x') < y' < H_+ \right\}$$

and

$$\left\{ \boldsymbol{r}' = (x', y') : x' = x + \frac{L}{2},\ f(x') < y' < H_+ \right\}$$

cancel each other out. The integral along

$$\left\{\boldsymbol{r}' = (x', y') : x - \frac{L}{2} < x' < x + \frac{L}{2},\ y' = H_+\right\}$$

can be shown [42] to equal 0, since the propagating waves of u are outgoing while those of v are incoming (since on this contour $y' > y$). Therefore, letting $H_+ \to \infty$, we have

$$\psi^{\text{scat}}(\boldsymbol{r}) = \int_{\mathcal{P}(x)} \left(\psi^{\text{scat}}(\boldsymbol{r}')\frac{\partial \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r}')} - \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')\frac{\partial \psi^{\text{scat}}(\boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r}')}\right) ds(\boldsymbol{r}'), \quad y > f(x). \tag{2.28}$$

Similarly, for some $H_- < \min[f(x)]$, we let

$$D_- = D_-(x, H_-) \equiv \left\{\boldsymbol{r}' = (x', y') : x - \frac{L}{2} < x' < x + \frac{L}{2},\ H_- < y' < f(x')\right\}, \tag{2.29}$$

$u(x', y') \equiv \psi^{\text{inc}}(\boldsymbol{r}')$ and $v(x', y') \equiv \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')$, but we keep $\boldsymbol{r} \in D_+$ as before. Here, both u and v solve the Helmholtz equation in D_-, so

$$\int_{D_-} \left\{u\left[\Delta v + k^2 v\right] - v\left[\Delta u + k^2 u\right]\right\} d\boldsymbol{r}' = 0. \tag{2.30}$$

Also,

$$\int_{D_-} \left[u\left(\Delta v + k^2 v\right) - v\left(\Delta u + k^2 u\right)\right] d\boldsymbol{r}' = \int_{\partial D_+} \left(u\frac{\partial v}{\partial \boldsymbol{\nu}(\boldsymbol{r}')} - v\frac{\partial u}{\partial \boldsymbol{\nu}(\boldsymbol{r}')}\right) ds(\boldsymbol{r}'), \tag{2.31}$$

where here we have defined $\boldsymbol{\nu}(\boldsymbol{r}')$ as the external normal to the boundary of D_- so that on $y = f(x)$ it is the same normal as before. Similar to the previous calculations, the integrals along

$$\left\{\boldsymbol{r}' = (x', y') : x' = x - \frac{L}{2},\ H_- < y' < f(x')\right\}$$

and

$$\left\{\boldsymbol{r}' = (x', y') : x' = x + \frac{L}{2},\ H_- < y' < f(x')\right\}$$

cancel each other out, while the integral along

$$\left\{\boldsymbol{r}' = (x', y') : x - \frac{L}{2} < x' < x + \frac{L}{2},\ y' = H_-\right\}$$

is 0. Thus,

$$0 = \int_{\mathcal{P}(x)} \left(\psi^{\text{inc}}(\boldsymbol{r}') \frac{\partial \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r}')} - \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}') \frac{\partial \psi^{\text{inc}}(\boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r}')} \right) ds(\boldsymbol{r}'), \quad y > f(x). \tag{2.32}$$

Adding (2.28) and (2.32)—using the TE/sound-soft boundary condition (2.7)—results in

$$\psi^{\text{scat}}(\boldsymbol{r}) = -\int_{\mathcal{P}(x)} \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}') \frac{\partial \psi(\boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r}')} \, ds(\boldsymbol{r}'), \quad y > f(x), \tag{2.33}$$

which may be re-written as

$$\psi(\boldsymbol{r}) + \int_{\mathcal{P}(x)} \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}') \frac{\partial \psi(\boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r}')} \, ds(\boldsymbol{r}') = \psi^{\text{inc}}(\boldsymbol{r}), \quad y > f(x), \tag{2.34}$$

where $\psi(\boldsymbol{r}) \equiv \psi^{\text{inc}}(\boldsymbol{r}) + \psi^{\text{scat}}(\boldsymbol{r})$ is the total field. Taking the normal derivative at the point $\boldsymbol{r}$ as a limit from above the surface $y = f(x)$, as was done for $\psi^{\text{scat}}(\boldsymbol{r})$ in (2.10), results in the second-kind integral equation

$$\frac{1}{2}\mu(\boldsymbol{r}) + \int_{\mathcal{P}(x)} \frac{\partial \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \, ds(\boldsymbol{r}') = \frac{\partial \psi^{\text{inc}}(\boldsymbol{r})}{\partial \boldsymbol{\nu}(\boldsymbol{r})}, \quad y = f(x), \tag{2.35}$$

where

$$\mu(\boldsymbol{r}) \equiv \frac{\partial \psi(\boldsymbol{r})}{\partial \boldsymbol{\nu}(\boldsymbol{r})}. \tag{2.36}$$

See, e.g., [19] for details concerning this limit process; the singularity of the periodic Green's function at $\boldsymbol{r} = \boldsymbol{r}'$ requires careful treatment.

Remark 2.1.6. *The formula (2.33) for $\psi^{scat}(\boldsymbol{r})$ is continuous in $\boldsymbol{r}$ at the interface $y = f(x)$ [19]. Thus, taking the limit as $y \to f(x)$ from above without applying the normal derivative results in the first-kind integral equation*

$$\int_{\mathcal{P}(x)} \Phi_{per}(\boldsymbol{r}, \boldsymbol{r}') \frac{\partial \psi(\boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r}')} \, ds(\boldsymbol{r}') = \psi^{inc}(\boldsymbol{r}), \quad y = f(x), \tag{2.37}$$

where we have again employed the boundary condition (2.7). This equation is ill-posed [35], and we do not make use of it in the method we present in this thesis.

For the TM/sound-hard system, we represent $\psi^{\text{scat}}(\boldsymbol{r})$ as

$$\psi^{\text{scat}}(\boldsymbol{r}) = \int_{\mathcal{P}(x)} \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}') \mu(\boldsymbol{r}') \, ds(\boldsymbol{r}'), \quad y > f(x), \tag{2.38}$$

for some unknown α quasi-periodic continuous function $\mu(\boldsymbol{r})$; this is similar to the TE/sound-soft case (2.33), and it can be verified as being a radiating solution of the two-dimensional Helmholtz equation. Applying the same normal derivative with respect to $\boldsymbol{r}$ as before and using the TM/sound-hard boundary condition (2.8), we have

$$\frac{1}{2}\mu(\boldsymbol{r}) - \int_{\mathcal{P}(x)} \frac{\partial \Phi_{\text{per}}(\boldsymbol{r},\boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r})}\mu(\boldsymbol{r}')\, ds(\boldsymbol{r}') = \frac{\partial \psi^{\text{inc}}(\boldsymbol{r})}{\partial \boldsymbol{\nu}(\boldsymbol{r})}, \quad y = f(x). \tag{2.39}$$

We note that the only difference between (2.35) and (2.39) is the sign in front of the integral.

Remark 2.1.7. *The function $\mu(\boldsymbol{r})$ in equations (2.35) and (2.39) is known as the "density" [20]. For these and other integral equations, it is related to—if not identical to—what is sometimes called the "surface current" [42, 56].*

2.1.4 Computation of Scattering Efficiencies

After computing $\mu(\boldsymbol{r})$ for either TE/sound-soft and TM/sound-hard scattering, the scattered field $\psi^{\text{scat}}(\boldsymbol{r})$ can be computed point-wise everywhere in $y > f(x)$ using (2.33) (for the TE/sound-soft case) or (2.38) (for the TM/sound-hard case). As discussed in Section 2.1.2, however, for $y > \max[f(x)]$ and for $k \neq \pm\alpha_n$ (k which are not Wood Anomaly values), $\psi^{\text{scat}}(\boldsymbol{r})$ can also be represented by the Rayleigh expansion (2.17) consisting of propagating and evanescent plane waves. In particular, $\psi^{\text{scat}}(\boldsymbol{r})$ can be approximated in the "far field region" (i.e., as $y \to \infty$) by the sum $\sum_{n\in U} B_n e^{i\alpha_n x + i\beta_n y}$ of the finitely many propagating waves—a useful approximation in many practical cases. Although the numerical method of this thesis focuses upon accurately and efficiently computing $\mu(\boldsymbol{r})$ and can easily be applied toward computing $\psi^{\text{scat}}(\boldsymbol{r})$ for all $y > f(x)$ (including $f(x) < y < \max[f(x)]$ if desired), we will use the Rayleigh expansion-based far field approximation throughout the remainder of this thesis.

To determine the expansion's coefficients B_n for the TE/sound-soft case, we substitute the periodic Green's function as given in (2.22) and the Rayleigh expansion (2.17)

into (2.33), which for $y > \max[f(x)]$ results in

$$\begin{aligned}
\psi^{\text{scat}}(\boldsymbol{r}) &= \sum_{n=-\infty}^{\infty} B_n e^{i\alpha_n x + i\beta_n y} \\
&= -\int_{\mathcal{P}(x)} \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')\mu(\boldsymbol{r}')\, ds(\boldsymbol{r}') \\
&= -\int_{x-\frac{L}{2}}^{x+\frac{L}{2}} \frac{i}{2L} \sum_{n=-\infty}^{\infty} \frac{e^{i\alpha_n(x-x')+i\beta_n[y-f(x')]}}{\beta_n} \mu(\boldsymbol{r}')\sqrt{1+[f'(x')]^2}\, dx' \\
&= \sum_{n=-\infty}^{\infty} e^{i\alpha_n x + i\beta_n y} \frac{1}{2iL\beta_n} \int_0^L \mu(\boldsymbol{r}') e^{-i\alpha_n x' - i\beta_n f(x')} \sqrt{1+[f'(x')]^2}\, dx',
\end{aligned} \tag{2.40}$$

where we have used the L-periodicity of the integrand and

$$ds(\boldsymbol{r}') = \sqrt{1+[f'(x')]^2}\, dx' \tag{2.41}$$

along $y' = f(x')$. Orthogonality of the functions $e^{i\alpha_n x + i\beta_n y} = e^{i\frac{2\pi n}{L}x} e^{i\alpha x + i\beta_n y}$ in x for $x \in [0, L]$ implies that for the TE/sound-soft case

$$B_n = \frac{1}{2iL\beta_n} \int_0^L \mu(\boldsymbol{r}') e^{-i\alpha_n x' - i\beta_n f(x')} \sqrt{1+[f'(x')]^2}\, dx'. \tag{2.42}$$

Similarly, it can be shown that for TM/sound-hard scattering

$$B_n = -\frac{1}{2iL\beta_n} \int_0^L \mu(\boldsymbol{r}') e^{-i\alpha_n x' - i\beta_n f(x')} \sqrt{1+[f'(x')]^2}\, dx'. \tag{2.43}$$

These formulas for the coefficients B_n are valid for all n, i.e., valid for the propagating and evanescent waves.

For the propagating waves ($n \in U$) which comprise the far field approximation of the solution, we define

$$e_n \equiv \frac{\beta_n}{\beta} |B_n|^2 \tag{2.44}$$

to be the "scattering efficiency." This quantity is the fraction of energy scattered in the nth direction of propagation, and it can be shown (see, e.g., [48]) that

$$\sum_{n \in U} e_n = 1. \tag{2.45}$$

The "energy balance criterion" (2.45) can be physically interpreted as a statement of conser-

vation of energy; the unit energy incident plane wave $\psi^{\text{inc}}(\boldsymbol{r}) = e^{i(k\boldsymbol{\alpha}\cdot\boldsymbol{r})}$ is perfectly reflected into the far field region along the finitely many directions $(\frac{\alpha_n}{k}, \frac{\beta_n}{k})$, $n \in U$ with energy fraction e_n in the nth direction.

Thus, one measurement of the accuracy of a computational method is the comparison of its computed scattering efficiencies to those previously published or previously computed by an alternate method. Furthermore, the energy balance criterion provides a measure of the error of the sum of computed efficiencies, which is particularly useful when studying the numerical parameter dependence of a computational method or investigating previously unstudied scattering systems. We make extensive use of such measurements in the numerical results we report in Chapter 4.

2.2 Periodic Green's Function vs. Partition of Unity

As discussed in Section 2.1, one approach to the scattering problem is to solve an integral equation. Equations (2.35) (for TE/sound-soft scattering) and (2.39) (for TM/sound-hard scattering) are examples of such an integral equation; other examples also have been derived and used [21, 42]. In all of these equations, the periodic Green's function $\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')$ (and/or a normal derivative, e.g., $\frac{\partial \Phi_{\text{per}}(\boldsymbol{r},\boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r})}$ on $y = f(x)$) appears in the integrals.

To solve such an integral equation, many numerical methods, e.g., most of those described in [13, 21], use one or more techniques to compute the periodic Green's function when generating a finite linear system approximation of the equation (special treatment is involved in the numerical quadrature at $\boldsymbol{r} = \boldsymbol{r}'$). These techniques use forms of $\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')$ different from the spatial form (2.21) in order to accelerate its computation (see, e.g., [39] for a discussion of many of these forms) and thus accelerate the generation of the approximating matrix equation for the scattering problem.

Our method, however, does not evaluate $\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')$. Instead, based upon recasting the integral in each of the scattering equations (2.35) and (2.39) as an improper integral over the infinite scattering surface, it computes a related quantity that has been multiplied by a C^∞ windowing function from a "partition of unity" (a set of such functions which sum to 1 throughout $x \in (-\infty, \infty)$; see, e.g., [14, 15]). The smooth decay of this windowing function to 0 results in a significantly more accurate approximation of the improper integral than would be the case if a rectangular window were used: for a rectangular window, the error

of approximation decreases as the inverse square root of the window width, while the error decreases super-algebraically in the window width when using the smooth window. Since a rectangular window corresponds to using a truncated series approximation of the spatial form of $\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')$, the recasting and smooth windowing are a way of dramatically accelerating the numerical quadrature of the integral. Thus, these steps, like the other methods' formulas for computing the periodic Green's function, together form a key component of our algorithm.

In addition to its facility in the approximation of the integral operator, the smooth windowing function we use in our algorithm also permits desirable convergence properties for our numerical quadrature method. Integrands multiplied by this function can be periodically extended (the period being larger than the length of the support of the window), allowing for the implementation of a spectrally accurate quadrature rule which, under certain conditions, is super-algebraically convergent. See Section 2.5 for a description of this method and Section 3.2 for a discussion of its properties.

We briefly review the forms of the periodic Green's function discussed in [4, 13, 21] (Section 2.2.1), since we will later compare computational results of our method to some of those from this literature. Then, for our algorithm, we introduce a partition of unity and a different expression of the integral in the scattering equations (Section 2.2.2). We motivate the usefulness of this reformulation by evaluating a simple example that is closely related to the computations performed by the numerical methods under consideration (Section 2.2.3).

2.2.1 Forms of the Periodic Green's Function

On $y = f(x)$ and $y' = f(x')$, the spatial form of the periodic Green's function—given by (2.21)—is

$$\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}') = \frac{i}{4} \sum_{n=-\infty}^{\infty} e^{-i\alpha nL} H_0^1(ku(x, x' - nL)), \tag{2.46}$$

where we have defined

$$\begin{aligned} u(x, x') &\equiv |\boldsymbol{r} - \boldsymbol{r}'|\big|_{\boldsymbol{r}=(x,f(x)),\quad \boldsymbol{r}'=(x',f(x'))} \\ &= \sqrt{(x - x')^2 + [f(x) - f(x')]^2} \end{aligned} \tag{2.47}$$

and have used the L-periodicity of $f(x)$. This form is mentioned in [4, 13, 21] due to its connection with the fundamental solution given by (2.18), but it is not used computationally

in the methods described in these papers due to its slow convergence [39]. We will make use of it when deriving our algorithm, however.

Another expression for the periodic Green's function is the spectral form (2.22) for $(x', y') \in \mathcal{D}$, where

$$\begin{aligned} \mathcal{D} &= \mathcal{D}(x, y) \\ &\equiv \left\{(x', y') \in \mathbb{R}^2 : (x - x', y - y') \neq (nL, 0) \text{ for any } n \in \mathbb{Z}\right\}. \end{aligned} \tag{2.48}$$

It has been shown [16] that in compact subsets of $\mathcal{D}\times\mathcal{K}$, where $\mathcal{K}$ is given by (2.24), the series converges uniformly but cannot be term-wise differentiated in y' at $y' = y$ (analogously, it cannot be term-wise differentiated in y at $y = y'$ in similar compact subsets of (x, y, k)). Since the convergence is rapid within these subsets, methods discussed in [13, 21] compute a truncated series approximation of (2.22) (and/or a normal derivative) on $y = f(x)$ and $y' = f(x')$ wherever the conditions x' is away from x and $f(x')$ is away from $f(x)$ both hold.

A second computationally advantageous form of $\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')$ is [13, 54]

$$\begin{aligned} \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}') &= \frac{i}{4} H_0^1(k\,|(x - x', y - y')|) \\ &+ \frac{1}{\pi} e^{ikL[1+\sin(\theta)]} e^{ik(x-x')} \int_0^\infty \frac{e^{-k(x'-x+L)u^2}}{1 - e^{ikL[1+\sin(\theta)]} e^{-kLu^2}} \frac{\cos\left[k(y - y')u(u^2 - 2i)^{\frac{1}{2}}\right]}{(u^2 - 2i)^{\frac{1}{2}}}\, du \\ &+ \frac{1}{\pi} e^{ikL[1-\sin(\theta)]} e^{ik(x'-x)} \int_0^\infty \frac{e^{-k(x-x'+L)u^2}}{1 - e^{ikL[1-\sin(\theta)]} e^{-kLu^2}} \frac{\cos\left[k(y - y')u(u^2 - 2i)^{\frac{1}{2}}\right]}{(u^2 - 2i)^{\frac{1}{2}}}\, du. \end{aligned} \tag{2.49}$$

The integrals, which exist for k which are not Wood Anomaly values, have integrands which decay exponentially in u, and they (and/or their normal derivatives) can be computed to a desired degree of accuracy on $y = f(x)$ and $y' = f(x')$ for x near x' (including at $x' = x$) by using appropriate numerical quadratures over sufficiently large subintervals. This is the approach taken in [13, 21]; approaches for the special treatment of the singularity of $\frac{i}{4} H_0^1(k\,|(x - x', f(x) - f(x'))|)$ at $x' = x$ are discussed in detail in these papers.

2.2.2 Recasting the Integral Equations and Applying a Partition of Unity

The forms of the periodic Green's function $\Phi_{\text{per}}(\boldsymbol{r},\boldsymbol{r}')$ given in (2.22) and (2.49), finite approximations of which are used by methods described in [13, 21] because of their computational efficiency, can be used to compute the improper integral

$$\int_{\mathcal{P}(x)} \frac{\partial \Phi_{\text{per}}(\boldsymbol{r},\boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}')\, ds(\boldsymbol{r}'), \quad y = f(x)$$

in (2.35) and (2.39). Our method treats this integral in a very different manner, however. To start with, it can be re-expressed as an improper integral over an infinite domain:

$$\begin{aligned}
&\int_{\mathcal{P}(x)} \frac{\partial \Phi_{\text{per}}(\boldsymbol{r},\boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}')\, ds(\boldsymbol{r}') \\
&\quad = \int_{x-\frac{L}{2}}^{x+\frac{L}{2}} \left(\frac{\partial}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \frac{i}{4} \sum_{n=-\infty}^{\infty} e^{-i\alpha n L} H_0^1(ku(x,x'-nL)) \right) \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx' \\
&\quad = \int_{-\infty}^{\infty} \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx' \\
&\quad = \int_{y'=f(x')} \frac{\partial \Phi(\boldsymbol{r},\boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx', \quad y = f(x),
\end{aligned} \tag{2.50}$$

where we have used the formulas (2.25), (2.46) and (2.41) for the domain $\mathcal{P}(x)$, the periodic Green's function $\Phi_{\text{per}}(\boldsymbol{r},\boldsymbol{r}')$ (spatial form) and the differential arc length $ds(\boldsymbol{r}')$, respectively, along with the L-periodicity of $f(x')$ and the α quasi-periodicity of $\mu(\boldsymbol{r}')$. Thus, with this recasting, (2.35) and (2.39) become

$$\frac{1}{2}\mu(\boldsymbol{r}) + \int_{-\infty}^{\infty} \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx' = \frac{\partial \psi^{\text{inc}}(\boldsymbol{r})}{\partial \boldsymbol{\nu}(\boldsymbol{r})}, \quad y = f(x), \tag{2.51}$$

and

$$\frac{1}{2}\mu(\boldsymbol{r}) - \int_{-\infty}^{\infty} \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx' = \frac{\partial \psi^{\text{inc}}(\boldsymbol{r})}{\partial \boldsymbol{\nu}(\boldsymbol{r})}, \quad y = f(x). \tag{2.52}$$

Now, the integral in these equations can be split into a sum of integrals by using a partition of unity. Letting

$$S(x,x_0,x_1) \equiv \begin{cases} 1 & , \quad |x| \le x_0 \\ \exp\left(\frac{2e^{-1/u}}{u-1}\right) & , \quad x_0 < |x| < x_1, \ u = \frac{|x|-x_0}{x_1-x_0} \\ 0 & , \quad |x| \ge x_1, \end{cases} \tag{2.53}$$

$$P_1(x,x',c,A) = S(x'-x, cA, A), \quad 0 < c < 1 \tag{2.54}$$

and

$$P_2(x,x',c,A) = 1 - P_1(x,x',c,A), \tag{2.55}$$

we have

$$\begin{aligned} &\int_{-\infty}^{\infty} \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx' \\ &\quad = \int_{-\infty}^{\infty} P_1(x,x',c,A) \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx' \\ &\quad + \int_{-\infty}^{\infty} P_2(x,x',c,A) \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx' \\ &\quad = \int_{x-A}^{x+A} P_1(x,x',c,A) \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx' \\ &\quad + \int_{-\infty}^{\infty} P_2(x,x',c,A) \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx'. \end{aligned} \tag{2.56}$$

Thus, through the application of the windowing function $P_1(x,x',c,A)$, another approximation of the improper integral of (2.51) and (2.52)—i.e., an approximation other than the one made by substituting a finite approximation of the periodic Green's function into the equivalent integral in (2.50)—is given by the relation

$$\begin{aligned} &\int_{-\infty}^{\infty} \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx' \\ &\quad \approx \int_{-\infty}^{\infty} P_1(x,x',c,A) \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx' \\ &\quad = \int_{x-A}^{x+A} P_1(x,x',c,A) \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx', \end{aligned} \tag{2.57}$$

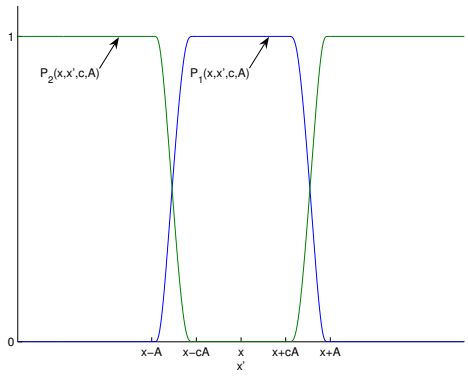


Figure 2.1: Partition of unity $P_1(x,x',c,A) + P_2(x,x',c,A) = 1$

and an alternative definition of this improper integral is

$$\begin{aligned}
&\int_{-\infty}^{\infty} \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx' \\
&\quad \equiv \lim_{A\to\infty} \int_{-\infty}^{\infty} P_1(x,x',c,A) \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx' \qquad (2.58) \\
&\quad = \lim_{A\to\infty} \int_{x-A}^{x+A} P_1(x,x',c,A) \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}\, dx'.
\end{aligned}$$

The approximation (2.57) is the one made by our algorithm.

We note that both $P_1(x,x',c,A)$ and $P_2(x,x',c,A)$ belong to $C^\infty(-\infty,\infty)$. Also, $\frac{\partial P_1}{\partial x'}(x,x',c,A) \to 0$ and $\frac{\partial P_2}{\partial x'}(x,x',c,A) \to 0$ for each fixed x as $A \to \infty$, and in the implementation of our method we choose c to be well away from 1. Thus, these functions comprising the partition of unity are both smooth and "gently-sloped"; these properties contribute (as we will illustrate) to accurately approximating the full integral. An example of the partition of unity is plotted in Figure 2.1.

2.2.3 Preliminary Simplified Case Study

To illustrate the effect of the partition of unity, we consider a simple example of an improper integral that contains some of the important features of the integral operator in (2.51) and (2.52) yet, unlike this operator, can be fully discussed with ease. The simplified integral is chosen on the basis of asymptotic considerations of the integrand

$$\frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1+[f'(x')]^2}$$

in these equations. A related—yet significantly more involved—analysis is given in Section 3.1.3 for the full integral operators of our method.

Since the integrand has a singularity at $x' = x$, we focus upon

$$\int_x^\infty \frac{i}{4}\frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})}\mu(\boldsymbol{r}')\sqrt{1+[f'(x')]^2}\,dx';$$

the integral

$$\int_{-\infty}^x \frac{i}{4}\frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})}\mu(\boldsymbol{r}')\sqrt{1+[f'(x')]^2}\,dx'$$

is analyzed similarly. Now,

$$\begin{aligned}
\frac{\partial \Phi(\boldsymbol{r},\boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r})} &= \frac{i}{4}\frac{\partial H_0^1(k|\boldsymbol{r}-\boldsymbol{r}'|)}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \\
&= \frac{i}{4}\,\nabla_{(x,y)} H_0^1(k|(x,y)-\boldsymbol{r}'|)\big|_{(x,y)=\boldsymbol{r}}\cdot\boldsymbol{\nu}(\boldsymbol{r}) \\
&= -\frac{i}{4}kH_1^1(ku(x,x'))\frac{(x-x',f(x)-f(x'))}{\sqrt{(x-x')^2+[f(x)-f(x')]^2}}\cdot\frac{(-f'(x),1)}{|(-f'(x),1)|} \\
&= \frac{i}{4}ku(x,x')H_1^1(ku(x,x'))\frac{f(x')-f(x)-(x'-x)f'(x)}{(x-x')^2+[f(x)-f(x')]^2}\frac{1}{\sqrt{1+[f'(x)]^2}}.
\end{aligned} \tag{2.59}$$

Because [3]

$$\lim_{z\to 0^+} zH_1^1(z) = \frac{2}{\pi i} \tag{2.60}$$

and

$$\begin{aligned}
&\lim_{x'\to x^+}\frac{f(x')-f(x)-(x'-x)f'(x)}{(x-x')^2+[f(x)-f(x')]^2} \\
&\quad= \lim_{x'\to x^+}\frac{f(x)+(x'-x)f'(x)+\frac{1}{2}(x'-x)^2f''(x)+\mathcal{O}((x'-x)^3)-f(x)-(x'-x)f'(x)}{(x-x')^2+[f(x)-f(x')]^2} \\
&\quad= \frac{f''(x)}{2+2\,[f'(x)]^2},
\end{aligned} \tag{2.61}$$

the integrand has a finite limit as $x' \to x^+$. On the other hand, the asymptotic expansion of $H_1^1(z)$ as $z\to\infty$ (z real and positive) is given by the formula [57]

$$H_1^1(z) \sim \left(\frac{2}{\pi z}\right)^{\frac{1}{2}} e^{i\left(z-\frac{3\pi}{4}\right)}\left[\sum_{m=0}^{p-1}\frac{(-1)^m\Gamma\left(\frac{3}{2}+m\right)}{m!\Gamma\left(\frac{3}{2}-m\right)(2iz)^m}+\mathcal{O}\left(z^{-p}\right)\right], \tag{2.62}$$

($\Gamma(z)$ is the well-known Gamma function) while

$$u(x,x') = \sqrt{(x-x')^2 + [f(x)-f(x')]^2} \sim x', \ x' \to \infty. \tag{2.63}$$

Also, the function $\sqrt{1+[f'(x')]^2}$ is L-periodic, so it can be represented by a Fourier series with basis functions of the form $e^{i\frac{2\pi n x'}{L}}$, $n \in \mathbb{Z}$, while $\mu(\boldsymbol{r}')$ being α quasi-periodic implies that it can be expanded in a series with basis functions of the form $e^{i\left(\alpha+\frac{2\pi n}{L}\right)x'}$, $n \in \mathbb{Z}$. So, as $x' \to \infty$, each mode of the integrand behaves like $\frac{e^{i\left(k+\alpha+\frac{2\pi n}{L}\right)x'}}{\sqrt{x'}}$ for some $n \in \mathbb{Z}$. Therefore, for purposes of illustration, we choose to analyze the improper integral

$$I_{\text{ex}}(\bar{k_n}) \equiv \int_0^\infty \frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\,dx', \tag{2.64}$$

where

$$\bar{k_n} \equiv k + \alpha + \frac{2\pi n}{L}, \quad n \in \mathbb{Z}, \tag{2.65}$$

since its integrand not only has the appropriate behavior as $x' \to \infty$ but also

$$\lim_{A\to\infty} \int_0^A \frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\,dx'$$

exists if $\bar{k_n} \neq 0$ (as we will show).

We begin by noting that for $\bar{k_n} = 0$ the integral

$$\int_0^\infty \frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\,dx' = \int_0^\infty \frac{1}{\sqrt{x'}}\,dx' \tag{2.66}$$

does not exist. $\bar{k_n} = k + \alpha + \frac{2\pi n}{L} = 0$ corresponds to k being a particular Wood Anomaly value (Remark 2.1.5); if k is not a Wood Anomaly value, then there is no $n \in \mathbb{Z}$ such that $\bar{k_n} = 0$. This is like the integral of the scattering equations (2.51) and (2.52): if k is a Wood Anomaly value, then that integral cannot be re-expressed in the form found in (2.35) and (2.39) (which use the periodic Green's function), and it also does not exist for most $\mu(\boldsymbol{r}')$. But, for $\bar{k_n} \neq 0$ we can compute a closed form expression for $I_{\text{ex}}(\bar{k_n})$. Letting

$$w \equiv \sqrt{\bar{k_n}x'} \longrightarrow dw = \sqrt{\bar{k_n}}\frac{1}{2\sqrt{x'}}dx' \tag{2.67}$$

(we may choose a branch of $\sqrt{\bar{k_n}}$ for negative $\bar{k_n}$ as well as for positive $\bar{k_n}$), the integral

becomes

$$\begin{aligned} I_{\text{ex}}(\bar{k_n}) &= \frac{2}{\sqrt{\bar{k_n}}} \int_0^\infty e^{iw^2}\, dw \\ &= \frac{2}{\sqrt{\bar{k_n}}} \int_0^\infty \left[\cos\left(w^2\right) + i \sin\left(w^2\right)\right]\, dw. \end{aligned} \tag{2.68}$$

Using complex analysis, or directly using the formulas [3]

$$\int_0^\infty \sin\left(w^2\right)\, dw = \int_0^\infty \cos\left(w^2\right)\, dw = \frac{1}{2}\sqrt{\frac{\pi}{2}}, \tag{2.69}$$

we determine that

$$I_{\text{ex}}(\bar{k_n}) = \sqrt{\frac{\pi}{2\bar{k_n}}} + i\sqrt{\frac{\pi}{2\bar{k_n}}}. \tag{2.70}$$

Remark 2.2.1. *For*

$$\int_{-\infty}^{x} \frac{i}{4} \frac{\partial H_0^1(ku(x,x'))}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \mu(\boldsymbol{r}') \sqrt{1 + [f'(x')]^2}\, dx',$$

an asymptotic analysis of the integrand shows that each of its modes behaves like $\frac{e^{i\left(-k+\alpha+\frac{2\pi n}{L}\right)x'}}{\sqrt{-x'}}$ *for some* $n \in \mathbb{Z}$ *as* $x' \to -\infty$*, so an appropriate simplified example is*

$$\int_{-\infty}^{0} \frac{e^{-i\bar{k_n}x'}}{\sqrt{-x'}}\, dx' = \int_0^\infty \frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\, dx', \tag{2.71}$$

where here

$$\bar{k_n} \equiv k - \alpha - \frac{2\pi n}{L},\ n \in \mathbb{Z}. \tag{2.72}$$

Again, $\bar{k_n} \neq 0$ *must hold for the integral to exist, and again this condition corresponds to* k *not being a particular Wood Anomaly value.*

In addition to this formula for $I_{\text{ex}}(\bar{k_n})$, we generate two different finite integral approximations which are of interest. Since

$$\begin{aligned} I_{\text{ex}}(\bar{k_n}) &= \int_0^A \frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\, dx' + \int_A^\infty \frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\, dx' \\ &= \int_0^\infty P_1(0,x',c,A) \frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\, dx' + \int_0^\infty P_2(0,x',c,A) \frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\, dx' \\ &= \int_0^A P_1(0,x',c,A) \frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\, dx' + \int_{cA}^\infty P_2(0,x',c,A) \frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\, dx', \end{aligned} \tag{2.73}$$

where $P_1(0,x',c,A)$ and $P_2(0,x',c,A)$ are determined by (2.53), (2.54) and (2.55), we let

$$I_{\text{per}}(\bar{k_n},A)=\int_0^A \frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\,dx' \tag{2.74}$$

and

$$I_{\text{part}}(\bar{k_n},c,A)\equiv\int_0^A P_1(0,x',c,A)\frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\,dx' \tag{2.75}$$

be two approximations for $I_{\text{ex}}(\bar{k_n})$, each of which converges to $I_{\text{ex}}(\bar{k_n})$ as $A\to\infty$. $I_{\text{part}}(\bar{k_n},c,A)$ can be viewed as the result of multiplying the kernel of the $I_{\text{ex}}(\bar{k_n})$ integral in (2.64) by a $C^\infty[0,\infty)$ windowing function equal to $P_1(0,x',c,A)$ for $x'\in[0,\infty)$, while $I_{\text{per}}(\bar{k_n},A)$ uses the rectangular (thus, discontinuous) windowing function which can be expressed as $\lim_{c\to 1}P_1(0,x',c,A)$ for $x'\in[0,\infty)$. Also, if $A=ML$ for some positive integer M, then $I_{\text{per}}(\bar{k_n},A)$ and $I_{\text{part}}(\bar{k_n},A)$ can be re-expressed as

$$\begin{aligned} I_{\text{per}}(\bar{k_n},A)&=\int_0^L\sum_{m=0}^{M-1}\frac{e^{i\bar{k_n}(x'+mL)}}{\sqrt{x'+\frac{2\pi}{\bar{k_n}}m}}\,dx' \\ &=\int_0^L\sum_{m=0}^{M-1}e^{i(k+\alpha)mL}\frac{e^{i\bar{k_n}x'}}{\sqrt{x'+\frac{2\pi}{\bar{k_n}}m}}\,dx' \end{aligned} \tag{2.76}$$

and

$$I_{\text{part}}(\bar{k_n},A)=\int_0^L\sum_{m=0}^{M-1}e^{i(k+\alpha)mL}P_1(0,x'+mL,c,A)\frac{e^{i\bar{k_n}x'}}{\sqrt{x'+\frac{2\pi}{\bar{k_n}}m}}\,dx'. \tag{2.77}$$

The integrand in the formula for $I_{\text{per}}(\bar{k_n},A)$ is similar to a mode of a truncated series approximation of the periodic Green's function's spatial form (2.21); the integrand in the formula for $I_{\text{part}}(\bar{k_n},A)$ is the same except that it has one or more of its terms multiplied by a smooth function with value between 0 and 1. Thus, for a fixed choice of c and increasing values of A we compare how quickly $I_{\text{per}}(\bar{k_n},A)$ and $I_{\text{part}}(\bar{k_n},c,A)$ approach $I_{\text{ex}}(\bar{k_n})$ by computing the approximation errors

$$\left|I_{\text{ex}}(\bar{k_n})-I_{\text{per}}(\bar{k_n},A)\right|=\left|\int_A^\infty\frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\,dx'\right| \tag{2.78}$$

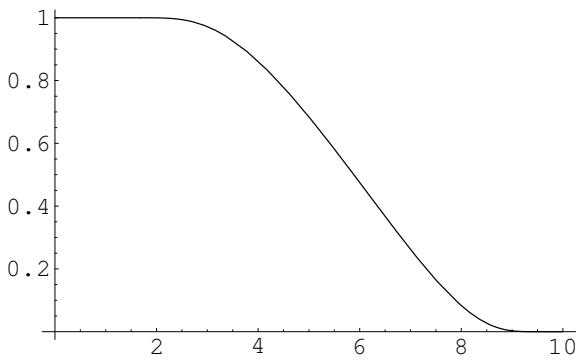


Figure 2.2: Partition of unity function $P_1(0, x', 0.1, 10)$ for $x' \in [0, 10]$

A	$\left\lvert I_{\text{ex}}(\bar{k_n}) - I_{\text{per}}(\bar{k_n}, A)\right\rvert$	$\left\lvert I_{\text{ex}}(\bar{k_n}) - I_{\text{part}}(\bar{k_n}, 0.1, A)\right\rvert$
10	5.0×10^{-2}	8.5×10^{-5}
20	3.6×10^{-2}	9.7×10^{-7}
25	3.2×10^{-2}	1.9×10^{-7}
50	2.3×10^{-2}	4.9×10^{-10}
75	1.8×10^{-2}	4.7×10^{-11}
100	1.6×10^{-2}	7.7×10^{-14}

Table 2.1: Approximation errors for various A ($\bar{k_n} = 2\pi$)

and

$$\left|I_{\text{ex}}(\bar{k_n}) - I_{\text{part}}(\bar{k_n}, c, A)\right| = \left|\int_{cA}^{\infty} P_2(0, x', c, A) \frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\, dx'\right|, \tag{2.79}$$

and in doing so we illustrate the impact of the smooth windowing function in our computations.

In particular, we consider $I_{\text{ex}}(2\pi) = \frac{1}{2} + \frac{1}{2}i$. Using Mathematica, we numerically evaluate the definite integrals for $I_{\text{per}}(2\pi, A)$ and $I_{\text{part}}(2\pi, c, A)$ using $c = 0.1$ and increasing values of A. A plot of one of the partition functions, $P_1(0, x', 0.1, 10)$, is given in Figure 2.2, and the approximation errors are given in Table 2.1. As A increases, $I_{\text{part}}(2\pi, 0.1, A)$ converges super-algebraically to $I_{\text{ex}}(2\pi)$ while $I_{\text{per}}(2\pi, A)$ converges very slowly, even though the limits of integration of their associated integrals is the same and $I_{\text{part}}(2\pi, 0.1, A)$ has the seemingly additional penalty of having its integrand multiplied by a function which is between 0 and 1 in part of the interval. Remarkably, $I_{\text{per}}(2\pi, 100)$ approximates $I_{\text{ex}}(2\pi)$ to about 2 digits while $I_{\text{part}}(2\pi, 0.1, 100)$ approximates $I_{\text{ex}}(2\pi)$ to nearly machine precision. We perform a similar study for $I_{\text{ex}}\left(\frac{\pi}{2}\right) = 1 + i$; the results are found in Table 2.2, and they follow a similar pattern to those for $I_{\text{ex}}(2\pi)$.

A	$\left\lvert I_{\text{ex}}(\bar{k_n}) - I_{\text{per}}(\bar{k_n}, A)\right\rvert$	$\left\lvert I_{\text{ex}}(\bar{k_n}) - I_{\text{part}}(\bar{k_n}, 0.1, A)\right\rvert$
100	6.4×10^{-2}	3.9×10^{-7}
200	4.5×10^{-2}	9.8×10^{-10}
300	3.7×10^{-2}	9.3×10^{-11}
400	3.2×10^{-2}	1.5×10^{-13}

Table 2.2: Approximation errors for various A ($\bar{k_n} = \frac{\pi}{2}$)

The differences in the errors of the approximations can be understood by computing asymptotic expansions of the integrals in (2.78) and (2.79) [5] (again, as we noted at the beginning of this section, the following analysis is related to, but much simpler than, the analysis for the actual integrals our method computes). Making a change of variables and using integration by parts,

$$\begin{aligned}
\left| \int_A^\infty \frac{e^{i\bar{k_n}x'}}{\sqrt{x'}}\, dx' \right| &= \left| \sqrt{A} \int_1^\infty \frac{e^{i\bar{k_n}A\bar{x'}}}{\sqrt{\bar{x'}}}\, d\bar{x'} \right| \\
&= \left| \sqrt{A}\, \frac{e^{i\bar{k_n}A\bar{x'}}}{i\bar{k_n}A\sqrt{\bar{x'}}} \right|_1^\infty + \frac{\sqrt{A}}{i\bar{k_n}A} \frac{1}{2} \int_1^\infty \frac{e^{i\bar{k_n}A\bar{x'}}}{\bar{x'}^{\frac{3}{2}}}\, d\bar{x'} \right| \\
&= \mathcal{O}\left(\frac{1}{\sqrt{\bar{k_n}}} \frac{1}{\sqrt{\bar{k_n}A}} \right).
\end{aligned} \tag{2.80}$$

This expansion indicates that as A increases from 10 to 100 there is about a factor of $\sqrt{10} \approx 3.16$ decrease in the size of the error in using the $I_{\text{per}}(\bar{k_n}, A)$ approximation found in (2.74), and such a decrease is observed in Table 2.1 (a factor of $\sqrt{4} \approx 2$ decrease is observed in Table 2.2 as A increases from 100 to 400). Stated another way, A must increase by a factor of 100, i.e., the number of terms in the series in (2.76) must increase by a factor of 100, in order to gain an additional digit of accuracy in approximation when using the rectangular window. Additionally, the expansion shows that the error is inversely proportional to the value of $\bar{k_n}$, i.e., it is inversely proportional to the square root of $\bar{k_n}$ multiplied by the square root of $\bar{k_n}A$; this is borne out by comparing the $A = 25$ value for $\bar{k_n} = 2\pi$ (Table 2.1) to the $A = 100$ value for $\bar{k_n} = \frac{\pi}{2}$ (Table 2.2), the $A = 50$ value for $\bar{k_n} = 2\pi$ to the $A = 200$ value for $\bar{k_n} = \frac{\pi}{2}$, etc. On the other hand,

$$P_2(0, x', c, A) = P_2\left(0, \frac{x'}{cA}, 1, \frac{1}{c}\right), \tag{2.81}$$

so that the error of approximation in using the smooth window as described in (2.75) is

$$\begin{aligned}\left|\int_{cA}^{\infty} P_2(0,x',c,A)\frac{e^{ik_n x'}}{\sqrt{x'}}\,dx'\right| &= \left|\sqrt{cA}\int_1^{\infty}\frac{P_2\left(0,\bar{x'},1,\frac{1}{c}\right)}{\sqrt{\bar{x'}}}e^{ik_n cA\bar{x'}}\,d\bar{x'}\right| \\ &= \left|\sqrt{cA}\,\frac{P_2\left(0,\bar{x'},1,\frac{1}{c}\right)}{\sqrt{\bar{x'}}}\frac{e^{ik_n A\bar{x'}}}{ik_n A}\right|_1^{\infty} \\ &\quad -\frac{\sqrt{cA}}{ik_n A}\int_1^{\infty}\frac{\partial}{\partial\bar{x'}}\left[\frac{P_2\left(0,\bar{x'},1,\frac{1}{c}\right)}{\sqrt{\bar{x'}}}\right]e^{ik_n cA\bar{x'}}\,d\bar{x'}\Bigg| \\ &= \frac{\sqrt{c}}{\sqrt{k_n}\sqrt{k_n A}}\left|\int_1^{\infty}\frac{\partial}{\partial\bar{x'}}\left[\frac{P_2\left(0,\bar{x'},1,\frac{1}{c}\right)}{\sqrt{\bar{x'}}}\right]e^{ik_n cA\bar{x'}}\,d\bar{x'}\right|.\end{aligned} \tag{2.82}$$

Integrating by parts p times—differentiating $\frac{P_2\left(0,\bar{x'},1,\frac{1}{c}\right)}{\sqrt{\bar{x'}}}$ (the derivatives of P_2 are non-zero only for $\bar{x'} \in \left(1,\frac{1}{c}\right)$) and integrating $e^{ik_n cA\bar{x'}}$ p times each—shows that the integral is $\mathcal{O}\left(\frac{1}{\sqrt{k_n}}\frac{1}{(k_n A)^{p-\frac{1}{2}}}\right)$ for every $p \geq 1$, since all of the boundary terms that arise from the integration by parts procedure are 0. Thus, the error of approximation decreases super-algebraically as A increases, and the relative sizes of the data between the Table 2.1 and Table 2.2 results for this approximation behaves as expected (e.g., the $A = 25$ value for $k_n = 2\pi$ is about half the size of the $A = 100$ value for $k_n = \frac{\pi}{2}$, etc.).

Remark 2.2.2. *The error depends upon c through the derivatives of P_2 (the magnitudes of these derivatives increase as c approaches 1) as well as the length of the interval $\bar{x'} \in \left(1,\frac{1}{c}\right)$ over which they are non-zero, so it is useful to choose a value of c that is non-zero but also well away from 1.*

Therefore, using a C^{∞} windowing function is decidedly superior to using a rectangular one (equivalent to using an truncation of the series for the periodic Green's function's spatial formula) when approximating the simple example integral (2.64). The same windowing also yields super-algebraic convergence in our algorithm's computations of the scattering problem; see Section 3.1.3 for a proof of this and Section 4.1 for numerical results that demonstrate this.

2.3 Two Representations of the Density

The TE/sound-soft integral equation (2.51) and the TM/sound-hard integral equation (2.52) are solved for the α quasi-periodic density $\mu(\boldsymbol{r})$, where $\alpha = k\sin(\theta)$ as stated in (2.13). These equations can be reformulated so as to be solved for an L-periodic function in the interval $x \in [0, L]$, where this function is equal to $\mu(\boldsymbol{r})$ divided by a known α quasi-periodic function. We derive two such sets of equations and unknowns in Sections 2.3.1 and 2.3.2. Then, in Section 2.3.3 we show how consideration of the types of waves induced by the incident wave impacting the grating $y = f(x)$ lead us to using one or the other of these sets in our computations depending upon the physical properties of the scattering system—in particular, depending upon whether multiple scattering is present. This use of two representations of $\mu(\boldsymbol{r})$ is another important component of our method, since for many scattering problems (especially high-frequency problems) significant computational savings can be achieved through a good choice of representation.

2.3.1 First Representation of the Density

Taking into account (2.59), the relation

$$\begin{aligned}\frac{\partial \psi^{\text{inc}}(\boldsymbol{r})}{\partial \boldsymbol{\nu}(\boldsymbol{r})} &= ik\boldsymbol{\alpha}\cdot\boldsymbol{\nu}(\boldsymbol{r})e^{ik\boldsymbol{\alpha}\cdot\boldsymbol{r}} \\ &= -ik\frac{\sin(\theta)f'(x)+\cos(\theta)}{\sqrt{1+[f'(x)]^2}}e^{ik[\sin(\theta)x-\cos(\theta)f(x)]}\end{aligned} \tag{2.83}$$

and dividing the TE/sound-soft equation (2.51) by $-ik\frac{e^{ik\sin(\theta)x}}{\sqrt{1+[f'(x)]^2}}$, we obtain

$$\mu_1(x) + \int_{-\infty}^{\infty} g(x,x')h(ku(x,x'))e^{-ik\sin(\theta)(x-x')}\mu_1(x')\,dx' = q_1(x), \quad 0 \le x \le L. \tag{2.84}$$

Here we have set

$$g(x,x') \equiv \frac{i}{2}\frac{f(x')-f(x)-(x'-x)f'(x)}{(x-x')^2+[f(x)-f(x')]^2}, \tag{2.85}$$

$$h(t) \equiv tH_1^1(t), \tag{2.86}$$

$$q_1(x) \equiv \left[\sin(\theta)f'(x)+\cos(\theta)\right]e^{-ik\cos(\theta)f(x)} \tag{2.87}$$

and

$$\mu_1(x) \equiv \mu(\boldsymbol{r}) \frac{e^{-ik\sin(\theta)x}}{-2ki} \sqrt{1 + [f'(x)]^2}. \tag{2.88}$$

Similarly, the TM/sound-hard equation (2.52) can be re-expressed as

$$\mu_1(x) - \int_{-\infty}^{\infty} g(x,x') h(ku(x,x')) e^{-ik\sin(\theta)(x-x')} \mu_1(x')\, dx' = q_1(x), \quad 0 \le x \le L. \tag{2.89}$$

In both the TE/sound-soft and TM/sound-hard cases, we have thus formulated the density $\mu(\boldsymbol{r})$ as the product of elementary functions, a function of the grating $y = f(x)$ and the L-periodic function $\mu_1(x)$—the unknown in the new equations. Due to their L-periodic nature, the equations (2.84) and (2.89) are specified as being solved for $x \in [0, L]$ for purposes of later discussion.

Given the e^{iz} factor in the asymptotic expansion (2.62) of $H_1^1(z)$ as $z \to \infty$, it is useful to re-express (2.84) and (2.89) in the form

$$\mu_1(x) \pm \int_{-\infty}^{\infty} g(x,x') \frac{h(ku(x,x'))}{e^{iku(x,x')}} e^{ik\phi_1(x,x')} \mu_1(x')\, dx' = q_1(x), \quad 0 \le x \le L, \tag{2.90}$$

where

$$\begin{aligned} \phi_1(x,x') &\equiv u(x,x') - \sin(\theta)(x-x') \\ &= \sqrt{(x-x')^2 + [f(x) - f(x')]^2} - \sin(\theta)(x-x'). \end{aligned} \tag{2.91}$$

This form of the integral, with the smooth windowing of Section 2.2.2 applied to it, will be analyzed in Section 3.1.3.

2.3.2 A Second Pair of Equations and an Alternate Unknown

Motivated by the work presented in [17], we re-express the scattering equations (2.51) and (2.52) in yet another form, which yields a new representation of $\mu(\boldsymbol{r})$. As we will show later, it is highly advantageous to use this alternate pair of equations for certain scattering configurations.

2.3.2.1 High-Frequency Asymptotic Series Ansatz and Its Physical Restrictions

In [17], a high-order high-frequency (i.e., large k) numerical method was developed for scattering problems under certain conditions. For TE/sound-soft scattering, the integral equation it solves is based upon a double-layer potential representation of the scattered field; stated in terms of the notation of this thesis, this equation reads

$$\frac{1}{2}\mu(\boldsymbol{r}) + \int_{y=f(x)} \frac{\partial \Phi(\boldsymbol{r},\boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r}')}\mu(\boldsymbol{r}')\,ds(\boldsymbol{r}') = -\psi^{\text{inc}}(\boldsymbol{r}), \quad y = f(x). \tag{2.92}$$

Now, dividing both sides of (2.92) by

$$\frac{1}{2}\psi^{\text{inc}}(\boldsymbol{r}) = \frac{1}{2}e^{i\alpha x - i\beta f(x)} = \frac{1}{2}e^{ik[\sin(\theta)x - \cos(\theta)f(x)]} \tag{2.93}$$

results in

$$e^{-i\alpha x + i\beta f(x)}\mu(\boldsymbol{r}) + 2\int_{y'=f(x')} \frac{\partial \Phi(\boldsymbol{r},\boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r}')}e^{-i\alpha x + i\beta f(x)}\mu(\boldsymbol{r}')\,ds(\boldsymbol{r}') = -2, \quad y = f(x), \tag{2.94}$$

and on the basis of this equation the method of [17] uses the asymptotic expression

$$\mu(\boldsymbol{r}) \sim e^{ik[\sin(\theta)x - \cos(\theta)f(x)]}\sum_{n=0}^{\infty}\frac{\nu_n(x)}{k^n} \tag{2.95}$$

as an ansatz: the unknown $\mu(\boldsymbol{r})$ is expressed as the product of a known, "rapidly oscillating" α quasi-periodic function (increasing oscillations as k increases) and an unknown, "slowly oscillating" L-periodic function (bounded number of oscillations as k increases) which can be represented asymptotically as $k \to \infty$ by a power series in $\frac{1}{k}$. Substituting this ansatz into (2.94), the L-periodic functions $\nu_n(x)$ are determined up to some chosen order (as k increases, the number of $\nu_n(x)$ that are needed for a given accuracy decreases) using Taylor/Fourier series expansions and analytic continuation to reinterpret certain divergent integrals. A similar ansatz was developed in [17] for TM/sound-hard scattering.

Many practical scattering systems, especially high-frequency systems, can be solved accurately and rapidly with this method. Also, unlike methods that use the periodic Green's function, this solver works for at least some Wood Anomaly values of k, as can be seen in the computational results given in [17] (some of these results are described in Appendix B).

However, the method requires that certain conditions are satisfied by the surface profile $y = f(x)$ and the incidence angle θ. In particular, $f(x)$ is presumed to be analytic (this allows certain Taylor series expansions to be used in the analysis). Also, given $t \equiv x' - x$, the condition

$$\frac{\partial}{\partial t}\sqrt{t^2 + [f(x+t) - f(x)]^2} > 0, \ t \geq 0 \tag{2.96}$$

is assumed to hold, which restricts the amplitude of $f(x)$ (an example given in [17] is that, for $f(x) = a\cos(x)$, the condition is satisfied for $a < 1$). Additionally, a joint restriction on $f(x)$ and θ is that the relation

$$\boldsymbol{\alpha} \cdot \boldsymbol{\nu}(\boldsymbol{r}) = -\frac{\sin(\theta)f'(x) + \cos(\theta)}{\sqrt{1 + [f'(x)]^2}} = 0 \longrightarrow f'(x) = -\cot(\theta) \tag{2.97}$$

is nowhere satisfied, where $\boldsymbol{\nu}(\boldsymbol{r})$ is the upward normal to the grating as defined in Section 2.1.1 (not to be confused with the functions $\nu_n(x)$). Physically, (2.97) describes a set of points $\boldsymbol{r} = (x, f(x))$ where rays of the incident wave are tangent to the surface. We call these locations "shadow boundaries" or "shadowing points," and we call the set of points where $\boldsymbol{\alpha} \cdot \boldsymbol{\nu}(\boldsymbol{r}) \geq 0$ (where the incident wave does not directly illuminate the grating) the "shadow region" [20]. So, the second restriction implies that there is no shadowing.

We also note that, as stated in [17], the function $\nu_0(x) = -2$ in the asymptotic series corresponds to the Kirchhoff approximation for the density $\mu(\boldsymbol{r})$. This high-frequency approximation, as is well known (see, e.g., [20]), results as the scattering surface at each point $\boldsymbol{r}$ is locally approximated by a tangent plane. Given the scattering equation (2.92)—in particular, the incident field

$$\begin{aligned} \psi^{\text{inc}}(\boldsymbol{r}) &= e^{i(k\boldsymbol{\alpha}\cdot\boldsymbol{r})} \\ &= e^{ik[\sin(\theta)x - \cos(\theta)f(x)]} \end{aligned} \tag{2.98}$$

at each point $\boldsymbol{r} = (x, f(x))$ on the grating—by the Law of Reflection and the planar approximation (for a plane wave reflecting from a planar surface, the angle of incidence equals the angle of reflection, where the angles are measured relative to the normal to the surface) the Kirchhoff approximation for the density is

$$\mu(\boldsymbol{r}) \approx -2e^{ik[\sin(\theta)x - \cos(\theta)f(x)]}. \tag{2.99}$$

This is equivalent to $\nu_0(x) = -2$ and $\nu_n(x) = 0,\ n = 1, 2, 3, \ldots$, and it also corresponds to approximating the integral in the equation as 0. The Kirchhoff approximation may be sufficient for certain applications, depending upon the size of k and the computational accuracy required, so we use it as a point of comparison for some of our results in Chapter 4.

Remark 2.3.1. *A different treatment for the shadowing that arises in the problem of scattering from a convex obstacle is discussed in [20]: the approximation $\mu(\boldsymbol{r}) \approx 0$ is used in the shadow region (i.e., where $\boldsymbol{\alpha} \cdot \boldsymbol{\nu}(\boldsymbol{r}) \geq 0$). The method of [17], however, does not deal with rough surface scattering cases that have shadowing (as stated earlier). We use the approximation of (2.99) for all points on the scattering surfaces—including those in the shadow regions that exist for certain cases—when applying the Kirchhoff approximation numerically in the computational results of Chapter 4.*

2.3.2.2 Second Representation of the Density

Because of the usefulness of the method presented in [17], particularly for large wavenumbers k, we establish a similar representation for the density in our scattering equations. We divide (2.51) and (2.52) by $-ik\frac{e^{ik[\sin(\theta)x - \cos(\theta)f(x)]}}{\sqrt{1+[f'(x)]^2}}$—equivalent to dividing (2.90) by $e^{-ik\cos(\theta)f(x)}$—to obtain

$$\mu_2(x) \pm \int_{-\infty}^{\infty} g(x, x') \frac{h(ku(x, x'))}{e^{iku(x,x')}} e^{ik\phi_2(x,x')} \mu_2(x')\, dx' = q_2(x), \quad 0 \leq x \leq L, \tag{2.100}$$

where

$$\begin{aligned} \mu_2(x) &\equiv \mu(\boldsymbol{r}) \frac{e^{-ik\sin(\theta)x + ik\cos(\theta)f(x)}}{-2ki} \sqrt{1 + [f'(x)]^2} \\ &= \mu(\boldsymbol{r}) \frac{e^{-ik\boldsymbol{\alpha}\cdot\boldsymbol{r}}}{-2ki} \sqrt{1 + [f'(x)]^2}\Big|_{\boldsymbol{r}=(x,f(x))} \\ &= \mu_1(x) e^{ik\cos(\theta)f(x)}, \end{aligned} \tag{2.101}$$

$$\begin{aligned} q_2(x) &\equiv \sin(\theta) f'(x) + \cos(\theta) \\ &= q_1(x) e^{ik\cos(\theta)f(x)} \end{aligned} \tag{2.102}$$

and

$$\begin{aligned}\phi_2(x,x') &\equiv u(x,x') - \sin(\theta)(x-x') + \cos(\theta)\left[f(x)-f(x')\right] \\ &= \left[|\boldsymbol{r}-\boldsymbol{r}'| - \boldsymbol{\alpha}\cdot(\boldsymbol{r}-\boldsymbol{r}')\right]_{\boldsymbol{r}=(x,f(x)),\quad \boldsymbol{r}'=(x',f(x'))} \\ &= \phi_1(x,x') + \cos(\theta)\left[f(x)-f(x')\right] \\ &= \sqrt{(x-x')^2+\left[f(x)-f(x')\right]^2} - (\sin(\theta),-\cos(\theta))\cdot\left(x-x', f(x)-f(x')\right).\end{aligned} \tag{2.103}$$

Thus, we have an alternate pair of equations—with unknown L-periodic function $\mu_2(x)$—which we may use to compute TE/sound-soft and TM/sound-hard scattering.

Unlike [17], in our method we do not assume that $\mu_2(x)$ is necessarily slowly oscillating in x with an asymptotic expansion of the form given in (2.95). Also, our method does not place any restrictions on the scattering configurations that can be examined; the grating can be very deep, and shadowing can occur.

When $\mu_2(x)$ is slowly oscillating, the coefficients of its Fourier series decay rapidly even for large k. Thus, sometimes significantly fewer Fourier coefficients are necessary to accurately represent $\mu_2(x)$ than are needed for $\mu_1(x)$, and it is useful to determine the configurations for which this is the case—as is done in the next section (Section 2.3.3). We briefly note here, though, that the ansatz (2.95) developed in [17] as well as the Kirchhoff approximation (2.99) each contain the rapidly oscillatory complex phase factor $e^{ik[\sin(\theta)x-\cos(\theta)f(x)]}$ multiplied by a slowly oscillating function (Section 2.3.2.1). Since this rapidly oscillating factor is equal to the incident field $\psi^{\text{inc}}(\boldsymbol{r})$ at $\boldsymbol{r}=(x,f(x))$ on the scattering surface, we expect that $\mu_2(x)$ will be slowly oscillating for cases in which no multiple reflections occur and each point on the grating is illuminated only by the incident field. This physical intuition is confirmed in the discussion found in the following section.

2.3.3 Physical Considerations in the Choice of Representations

The grating profile's height h ($h \equiv \max_{x,x'}|f(x)-f(x')|$), period L and shape (e.g., sinusoidal), together with the incidence angle θ and wavenumber k of the incident plane wave (assuming either TE/sound-soft or TM/sound-hard scattering), characterize the scattering systems under consideration in this thesis, and the interactions of the plane wave with the grating give rise to a number of types of scattering phenomena in such systems. We discuss

the types of scattering that can occur and relate these physical phenomena to the integral equations that we solve. By doing this, and by examining some example cases, we designate the set of scattering configurations for which we use (2.90) (computing $\mu_1(x)$) and the set for which we use (2.100) (computing $\mu_2(x)$), because making this designation leads to substantially more efficient computations in many cases; the representation $\mu_2(x)$ (2.101) is used for cases in which no multiple scattering is present, while $\mu_1(x)$ (2.88) is used for cases in which such scattering occurs.

2.3.3.1 Types of Scattering

Many kinds of scattering phenomena may arise when plane waves impinge upon the sort of gratings we are considering [7]. We call one such kind "simple reflections" or "single scattering." For this type of scattering, a ray of the incident wave impacts the grating at a point and then reflects back up to infinity at an angle determined by the Law of Reflection. Another kind is "multiple reflections" or "multiple scattering," in which a ray impacts the surface at one point and then impacts one or more other points on it before traveling back up to infinity. Finally, there is shadowing (introduced in Section 2.3.2.1); according to the Geometrical Theory of Diffraction, a ray impacting the grating tangentially at a shadowing point generates "creeping waves" that propagate along the scattering surface and re-radiate rays tangentially to the surface as they propagate [26, 33].

Remark 2.3.2. *For surfaces with corners and edges, which are not treated in this thesis but may be considered in future work, other types of scattering may occur; see [7] for details.*

For a given grating and incident field, one, some or even all of these kinds of scattering may occur. If the grating is sufficiently shallow (i.e., $\frac{h}{L}$ is small enough) and θ is sufficiently close to 0 (i.e., close enough to normal incidence), only simple reflections occur. If the grating is relatively deeper or the magnitude of the incidence angle is larger, however, then the other types of scattering may also occur. It is possible for multiple reflections to exist without shadowing, e.g., given a sufficiently deep grating and $\theta = 0$. But, it is not possible for shadowing to occur without the existence of multiple reflections, since the grating is at least twice continuously differentiable; the rays which initially impact the grating sufficiently near a shadowing point impinge upon the grating a second time at points near where the line tangent to the grating at that shadowing point intersects the grating a second time.

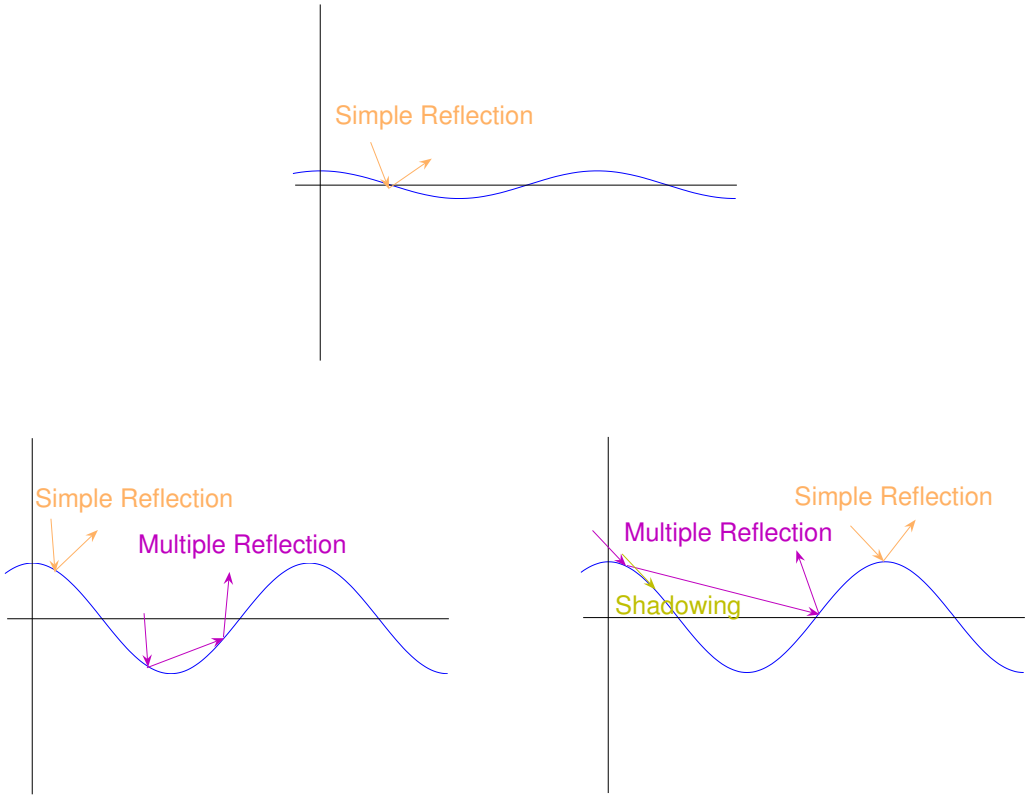


Figure 2.3: Case with only simple reflections (top), case with simple and multiple reflections (bottom left) and case with simple reflections, multiple reflections and shadowing (bottom right)

See Figure 2.3 for illustrations of these cases.

Remark 2.3.3. *We denote cases in which only simple reflections occur as "simple-reflection cases." Cases in which multiple reflections arise (with or without shadowing) are called "multiple-reflection cases."*

Using ray tracing, we can determine which scattering phenomena exist for any system we wish to consider. Instead, however, we develop certain numerical tests which are closely related to the functions found in the integral equations in (2.90) and (2.100). Such tests can be applied not only to individual systems, but also to whole classes of systems; in particular, we apply them to scattering from sinusoidal gratings of the form $f(x) = \frac{h}{2}\cos\left(\frac{2\pi x}{L}\right)$, and we make extensive use of the results throughout the remainder of this thesis.

2.3.3.2 Test for Multiple Reflections

For the equations in (2.100), which have the unknown $\mu_2(x)$, the kernel of the integral contains the phase function $\phi_2(x, x')$. Using this function, we prove a theorem that provides us with one numerical test for determining the types of scattering that exist for a given incident wave and grating profile. In particular, this test determines if multiple reflections (and possibly shadowing) are occurring or if there are only simple reflections.

Theorem 2.3.1. *For $\phi_2(x, x')$ (2.103), we have*

$$\frac{\partial \phi_2(x, x')}{\partial x'} = 0 \tag{2.104}$$

for some x and x' if and only if there are multiple reflections in the scattered field.

Proof. We have

$$\begin{aligned}
\frac{\partial \phi_2(x, x')}{\partial x'} &= \frac{\partial}{\partial x'} \left\{ \sqrt{(x - x')^2 + [f(x) - f(x')]^2} - \boldsymbol{\alpha} \cdot \big(x - x', f(x) - f(x')\big) \right\} \\
&= -\frac{x - x' + [f(x) - f(x')]\, f'(x')}{\sqrt{(x - x')^2 + [f(x) - f(x')]^2}} + \boldsymbol{\alpha} \cdot \big(1, f'(x')\big) \\
&= -\frac{(x - x', f(x) - f(x'))}{|(x - x', f(x) - f(x'))|} \cdot \big(1, f'(x')\big) + \boldsymbol{\alpha} \cdot \big(1, f'(x')\big)\,.
\end{aligned} \tag{2.105}$$

Since

$$\big|\big(1, f'(x')\big)\big| = \sqrt{1 + [f'(x')]^2} \neq 0, \tag{2.106}$$

it follows that $\left.\frac{\partial \phi_2(x,x')}{\partial x'}\right|_{x'=x_c} = 0$ if and only if

$$\frac{(x - x_c, f(x) - f(x_c))}{|(x - x_c, f(x) - f(x_c))|} \cdot \frac{(1, f'(x_c))}{|(1, f'(x_c))|} = \boldsymbol{\alpha} \cdot \frac{(1, f'(x_c))}{|(1, f'(x_c))|}. \tag{2.107}$$

Defining the unit vectors

$$\boldsymbol{d} \equiv \frac{(x - x', f(x) - f(x'))}{|(x - x', f(x) - f(x'))|} \tag{2.108}$$

and

$$\boldsymbol{\tau} \equiv \frac{(1, f'(x'))}{|(1, f'(x'))|}, \tag{2.109}$$

we re-express (2.107) as

$$[\boldsymbol{d}\cdot\boldsymbol{\tau}]_{x'=x_c} = \boldsymbol{\alpha}\cdot\boldsymbol{\tau}|_{x'=x_c}\,, \tag{2.110}$$

Geometrically, this equation tells us that at $x' = x_c$ the angle between the vectors $\boldsymbol{d}$ and $\boldsymbol{\tau}$ is the same as that between $\boldsymbol{\alpha}$ and $\boldsymbol{\tau}$. Noting that $\boldsymbol{d}|_{x'=x_c}$ is the unit vector pointing from $(x_c, f(x_c))$ to $(x, f(x))$, $\boldsymbol{\tau}|_{x'=x_c}$ is the tangent to the grating at $x' = x_c$ and $\boldsymbol{\alpha} = (\sin(\theta), -\cos(\theta))$ is the direction of propagation of the incident wave (Section 2.1.1), we see that equation (2.110) admits two types of solutions, namely:

1. The "shadowing" solutions

$$\boldsymbol{d}|_{x'=x_c} = \boldsymbol{\alpha} \tag{2.111}$$

   depicted in Figure 2.4, and

2. The "multiple reflection" solutions that arise as a ray of the incident wave reflects from the point $(x_c, f(x_c))$ onto either $(x, f(x))$ or a point in between (in accordance with the Law of Reflection). See Figure 2.5 for the $x_c > x$ case.

For solutions of the first type, consider the case $x_c > x$ (the $x_c < x$ case is handled similarly). By the Mean Value Theorem, there is a point $(\eta, f(\eta))$ for $\eta \in (x, x_c)$ at which

$$-\boldsymbol{\tau}|_{x'=\eta} = \boldsymbol{\alpha}, \tag{2.112}$$

i.e., $(\eta, f(\eta))$ is a shadowing point. The $x_c = x$ case is a degenerate version of the $x_c > x$ case, with $(x, f(x))$ being a shadowing point. See Figure 2.4 for illustrations of these cases. Thus, there are multiple reflections, particularly of the rays which initially impact the grating near the shadowing point (Section 2.3.3.1). Therefore, $\frac{\partial\phi_2(x,x')}{\partial x'}\Big|_{x'=x_c} = 0$ implies that there are multiple reflections.

Conversely, if there are multiple reflections, then there exist values x_1 and x_2 $(x_1 \neq x_2)$ such that a ray of the incident wave initially impinges the grating at $(x_2, f(x_2))$ and then reflects onto the grating at $(x_1, f(x_1))$. Since the ray obeys the Law of Reflection, this implies that $\frac{\partial\phi_2(x_1,x')}{\partial x'}\Big|_{x'=x_2} = 0$. □

Corollary 2.3.1. *If there are no multiple reflections in the scattered field, then*

$$\frac{\partial\phi(x,x')}{\partial x'} > 0, \quad x' > x \tag{2.113}$$

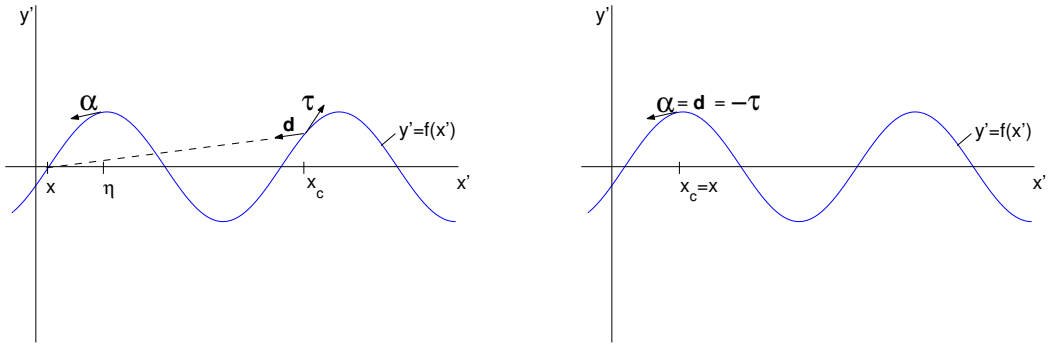


Figure 2.4: $x_c > x$ (left) and $x_c = x$ (right) shadowing cases

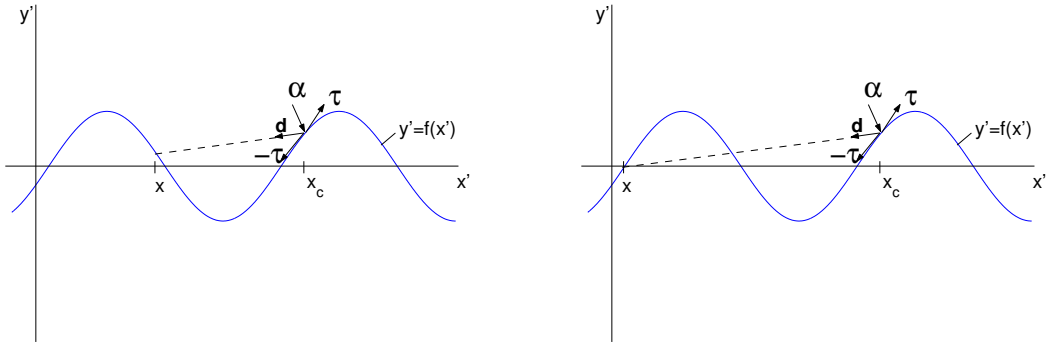


Figure 2.5: Instances of multiple reflections for $x_c > x$

and

$$\frac{\partial \phi(x, x')}{\partial x'} < 0, \quad x' < x. \tag{2.114}$$

Proof. For $x' \neq x$, we may write

$$\phi_2(x, x') = \left|\left(x - x', f(x) - f(x')\right)\right| (1 - \boldsymbol{\alpha} \cdot \boldsymbol{d}). \tag{2.115}$$

If there is no shadowing, then $\boldsymbol{\alpha} \cdot \boldsymbol{d} < 1$ for all x, x'. Therefore, for cases in which there are no multiple reflections (and thus no shadowing), $\phi_2(x, x') = 0$ for $x' = x$ while $\phi_2(x, x') > 0$ for $x' \neq x$, and the result follows by Theorem 2.3.1. □

2.3.3.3 Test for Shadowing

As stated in Section 2.3.2.1, shadowing occurs if there are points $\boldsymbol{r} = (x, f(x))$ such that $\boldsymbol{\alpha} \cdot \boldsymbol{\nu}(\boldsymbol{r}) = 0$, i.e., $f'(x) = -\cot(\theta)$. The converse also holds: if $(x, f(x))$ is a shadow point, then $f'(x) = -\cot(\theta)$. Thus, we have the test

$$f'(x) = -\cot(\theta) \tag{2.116}$$

for some x if and only if there is shadowing.

We note that the right-hand sides of the scattering equations in (2.90) and (2.100) are $q_1(x) = [\sin(\theta) f'(x) + \cos(\theta)]\, e^{-ik\cos(\theta) f(x)}$ and $q_2(x) = \sin(\theta) f'(x) + \cos(\theta)$, respectively. By (2.116), these functions vanish at the shadowing points. Thus, like $\phi_2(x, x')$, they are functions explicit in the integral equations we are solving which can be straightforwardly analyzed to test for the types of scattering inherent in a given system.

2.3.3.4 Height-to-Period Ratio vs. Incidence Angle

One implication of the analysis of the previous sections is that for any given grating profile we can determine the values of θ for which only simple reflections are induced by the incident wave, values for which multiple reflections also occur but shadowing does not and values for which both multiple reflections and shadowing arise. For shallow gratings, a large subinterval of $\theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ may satisfy the criterion of Section 2.3.3.2 for no multiple reflections existing, while there may be no such values for deep gratings. The test described in Section 2.3.3.3 indicates, however, that for every rough surface of the type considered in

this thesis there are always values of θ for which no shadowing occurs.

Another implication of the analysis is that for a set of scattering configurations with θ fixed and a variety of profiles we can determine which scattering phenomena are present in each case. In particular, if the set of profiles consists of one form that is scaled to various periods L or various heights h, we can generate functions of θ that define regions of (L, θ)-space or (h, θ)-space corresponding to cases with shadowing, with no shadowing but with multiple reflections, etc.

As an example (one that will be very useful later in this thesis), let us consider profiles of the form

$$f(x) = \frac{h}{2}\cos(2\pi x), \quad h > 0 \tag{2.117}$$

(height-to-period ratio $\frac{h}{L} = h$). For each value of θ, we can determine the minimum value of h for which a plane wave with incidence angle θ multiply reflects off of the grating—a value we denote as $h_{mult}(\theta)$. Only simple reflections arise in such a case if $h < h_{mult}(\theta)$. So, when examining the scattering from many gratings of this form, as we do in Chapter 4, we can refer to $h_{mult}(\theta)$ to determine if multiple scattering is present. A similar function with regard to shadowing also can be generated for these profiles.

To determine $h_{mult}(\theta)$ for $\theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, we derive and solve three equations of three unknowns: h, x_1 and x_2. The first equation is

$$\begin{aligned}
\frac{\partial \phi_2(x_1, x_2)}{\partial x_2} &= \frac{(x_2 - x_1, f(x_2) - f(x_1))}{|(x_2 - x_1, f(x_2) - f(x_1))|} \cdot (1, f'(x_2)) + (\sin(\theta), -\cos(\theta)) \cdot \left(1, f'(x_2)\right) \\
&= \frac{x_2 - x_1 + [f(x_2) - f(x_1)]\, f'(x_2)}{\sqrt{(x_2 - x_1)^2 + [f(x_2) - f(x_1)]^2}} + \sin(\theta) - f'(x_2)\cos(\theta) \\
&= 0,
\end{aligned} \tag{2.118}$$

which was shown in Section 2.3.3.2 to hold if and only if multiple reflections occur. For a given x_1 and h, there may be many solutions x_2 to this equation (Figure 2.6).

Now, given x_1, there is a minimum value of h for which equation (2.118) holds for some x_2. This is because

$$\lim_{h \to 0} \frac{\partial \phi_2(x_1, x_2)}{\partial x_2} = \begin{cases} 1 + \sin(\theta) & , \quad x_2 > x_1 \\ -1 + \sin(\theta) & , \quad x_2 < x_1, \end{cases} \tag{2.119}$$

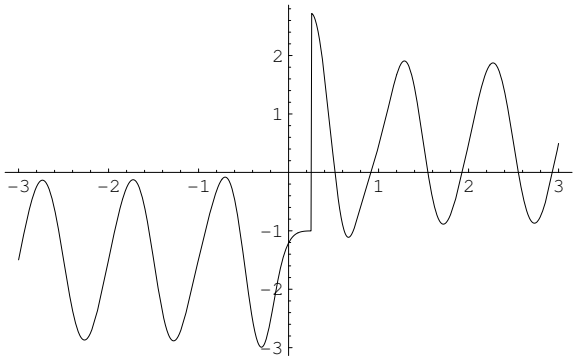


Figure 2.6: Plot of $\frac{\partial \phi_2(x_1,x_2)}{\partial x_2}$ with $\theta = -\frac{\pi}{6}$, $h = 0.5$ and $x_1 = 0.25$

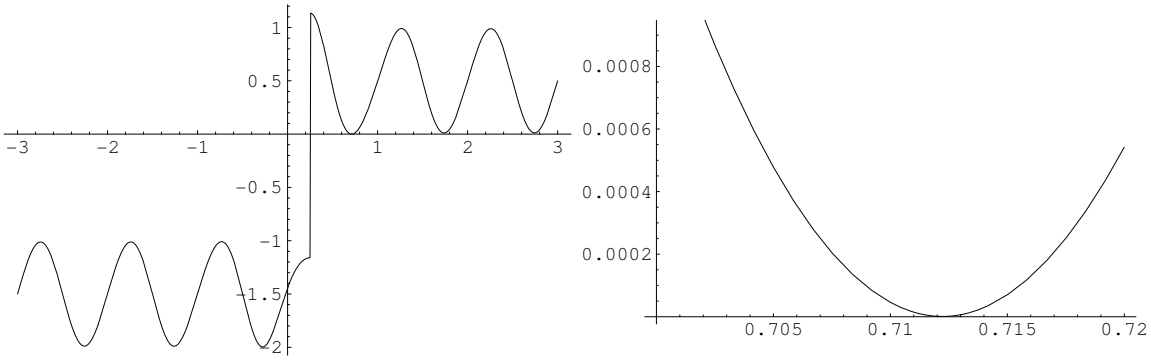


Figure 2.7: Plot of $\frac{\partial \phi_2(x_1,x_2)}{\partial x_2}$ with $\theta = -\frac{\pi}{6}$, $h \approx 0.179$ and $x_1 = 0.25$

where $1 + \sin(\theta) > 0$ and $-1 + \sin(\theta) < 0$ for $\theta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ (i.e., no solutions x_2 of the equation exist if h is sufficiently small), while $\frac{\partial \phi_2(x_1,x_2)}{\partial x_2} = 0$ for multiple values of x_2 if h is sufficiently large. For these values of h, x_1 and x_2, the second equation

$$\begin{aligned} \frac{\partial^2 \phi_2(x_1,x_2)}{\partial x_2^2} &= \frac{1 + [f'(x_2)]^2 + [f(x_2) - f(x_1)] f''(x_2)}{\sqrt{(x_2 - x_1)^2 + [f(x_2) - f(x_1)]^2}} \\ &\quad - \frac{\{x_2 - x_1 + [f(x_2) - f(x_1)] f'(x_2)\}^2}{\left\{\sqrt{(x_2 - x_1)^2 + [f(x_2) - f(x_1)]^2}\right\}^3} - f''(x_2)\cos(\theta) \\ &= 0 \end{aligned} \tag{2.120}$$

holds since $\frac{\partial \phi_2(x_1,x_2)}{\partial x_2}$ both equals 0 and has a local maximum or minimum at this value of x_2. This is illustrated in Figure 2.7.

Finally, for h, x_1 and x_2 which solve the first and second equations, the corresponding ray reflects from the grating at $(x_2, f(x_2))$ and re-impinges onto the grating at $(x_1, f(x_1))$

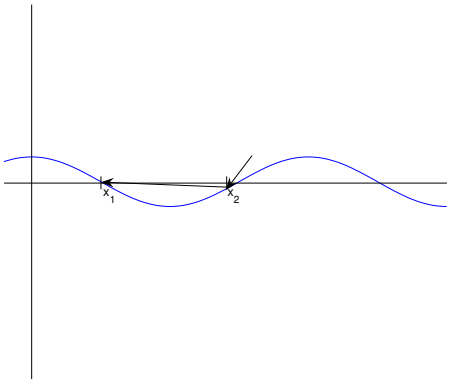


Figure 2.8: Plot of a ray and the grating profile with $\theta = -\frac{\pi}{6}$, $h \approx 0.179$, $x_1 = 0.25$ and $x_2 \approx 0.712$

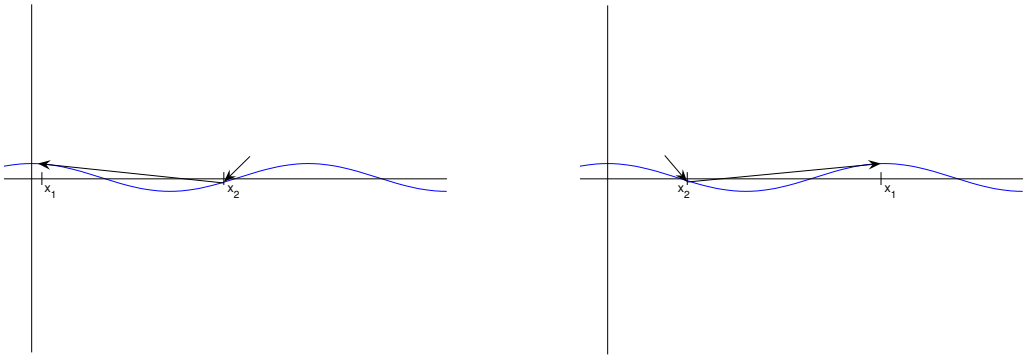


Figure 2.9: Plots of rays and the grating profile with $\theta < 0$ (left) and $\theta > 0$ (right)

(Figure 2.8). The third equation,

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(x_1), \tag{2.121}$$

arises from a physical consideration of the direction of this ray. For $h = h_{mult}(\theta)$, x_1 and x_2 are such that the ray impacts the scattering surface at $(x_1, f(x_1))$ tangentially (Figure 2.9), since if the impact were not tangential we could choose a grating with a smaller value of h such that the ray would still multiply reflect from it. Thus, the slope of the segment connecting $(x_2, f(x_2))$ to $(x_1, f(x_1))$ must equal the slope of the tangent to the surface at $(x_1, f(x_1))$.

If (2.121) holds, then (2.118) can be re-written somewhat more simply as

$$\frac{1+f'(x_1)f'(x_2)}{\operatorname{sgn}(x_2-x_1)\sqrt{1+[f'(x_1)]^2}}+\sin(\theta)-f'(x_2)\cos(\theta)=0, \tag{2.122}$$

and (2.120) becomes

$$\begin{aligned}0=&\frac{1+[f'(x_2)]^2+[f(x_2)-f(x_1)]f''(x_2)}{|x_2-x_1|\sqrt{1+[f'(x_1)]^2}}\\&-\frac{\{1+f'(x_1)f'(x_2)\}^2}{|x_2-x_1|\left\{\sqrt{1+[f'(x_1)]^2}\right\}^3}-f''(x_2)\cos(\theta).\end{aligned} \tag{2.123}$$

The three equations (2.121), (2.122) and (2.123) comprise the system that we solve.

This system of equations can be solved numerically for $\theta\in\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$. Good initial guesses for h, x_1 and x_2 are required, because there is not necessarily a unique solution; constraints on the guesses include $|x_1-x_2|<1$ (i.e., the two values are less than one period apart) as well as $x_1<x_2$ for $\theta<0$ and $x_1>x_2$ for $\theta>0$ (see Figure 2.9). Taking advantage of the inherent physical symmetry about $\theta=0$ and assuming small changes in the values of the solution given small changes in θ, we generate the function $h_{mult}(\theta)$ using MATLAB.

The minimum values of h for which shadowing occurs are considerably easier to compute. We substitute the formula

$$f'(x)=-\pi h\sin(2\pi x) \tag{2.124}$$

for the derivative of the surface profile into the shadowing equation (2.116) of Section 2.3.3.3 to obtain the relation

$$-\pi h\sin(2\pi x)=-\cot(\theta). \tag{2.125}$$

Since $0\le|\sin(2\pi x)|\le 1$, it follows that shadowing occurs for a particular value of θ if and only if

$$h\ge\frac{1}{\pi}\left|\cot(\theta)\right|. \tag{2.126}$$

Therefore, the minimum values of h for which shadowing occurs are given by the function

$$h_{shad}(\theta)\equiv\frac{1}{\pi}\left|\cot(\theta)\right|. \tag{2.127}$$

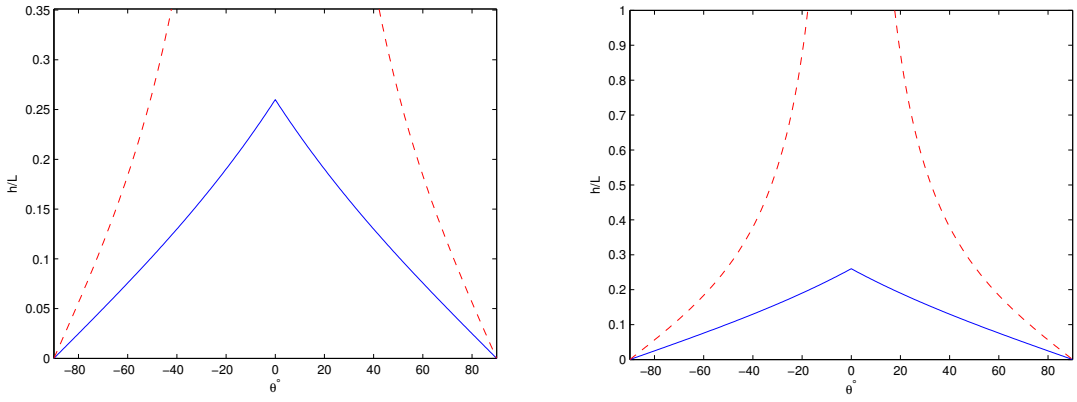


Figure 2.10: Plots of the multiple reflection threshold $h_{mult}(\theta)$ (solid line) and the shadowing threshold $h_{shad}(\theta)$ (dashed line) for the grating profile $f(x) = \frac{h}{2}\cos\left(\frac{2\pi x}{L}\right)$

The above results are straightforwardly extended to gratings of the form

$$f(x) = \frac{h}{2}\cos\left(\frac{2\pi x}{L}\right) \tag{2.128}$$

(height-to-period ratio $\frac{h}{L}$). As can be seen by making the changes of variables $\bar{x} \equiv \frac{x}{L}$, $\bar{y} \equiv \frac{y}{L}$ and $\bar{h} \equiv \frac{h}{L}$, there are only simple reflections of an incoming wave with incidence angle θ if $\frac{h}{L} < h_{mult}(\theta)$, and there are multiple reflections but no shadowing if $h_{mult}(\theta) \leq \frac{h}{L} < h_{shad}(\theta)$.

Plots of $h_{mult}(\theta)$ and $h_{shad}(\theta)$ are given in Figure 2.10. In addition to the symmetry of both functions about $\theta = 0$, we note that $h_{shad}(\theta) \to \infty$ as $\theta \to 0$ and that both $h_{mult}(\theta) \to 0$ and $h_{shad}(\theta) \to 0$ as $\theta \to \pm\frac{\pi}{2}$—also physically intuitive results. We will make extensive use of these plots when presenting our computational results in Chapter 4.

Remark 2.3.4. *Of course, more complicated grating forms can be examined in the above manner as well. For example, we can determine threshold functions $a_{mult}(\theta)$ and $a_{shad}(\theta)$ for multi-scale surfaces such as*

$$f(x) = \frac{a}{2}\left[\cos(2\pi x) + 0.04\sin(50\pi x)\right], \quad a > 0 \tag{2.129}$$

(in this particular case, there is the minor additional problem of numerically computing the

case	h	$\frac{k}{2\pi}$	θ
1	0.025	1.5	$\frac{\pi}{6}$
2	0.025	100.5	$\frac{\pi}{6}$
3	1.0	1.5	0
4	1.0	100.5	0
5	1.0	1.5	$\frac{\pi}{6}$
6	1.0	100.5	$\frac{\pi}{6}$

Table 2.3: Physical quantities for the examples of this section

maximum value of $|f'(x)|$ *for the* $a_{shad}(\theta)$ *formula). Certain gratings, e.g.,*

$$f(x) = 0.1\cos(2\pi x) + \frac{a}{2}\sin(50\pi x), \quad a > 0, \tag{2.130}$$

require more involved analysis, however, since for some incidence angles there are multiple reflections and perhaps also shadowing for all values of $a > 0$.

2.3.3.5 Examples Illustrating the Behavior of $\mu_1(x)$ and $\mu_2(x)$

As examples, we compute $\mu_1(x)$ and $\mu_2(x)$ for various typical TE/sound-soft scattering configurations. Using a grating profile of the form $f(x) = \frac{h}{2}\cos(2\pi x)$, we vary the grating height h as well as the wavenumber k and the incidence angle θ of the incident plane wave so as to observe the effects of the three types of scattering. Table 2.3 lists the cases we examine. Cases 1–2 only have simple reflections, Cases 3–4 have simple and multiple reflections but no shadowing and Cases 5–6 have all three types (Figure 2.10). The wavenumbers $k = 1.5 \times 2\pi,\ 100.5 \times 2\pi$ are chosen so as to be well away from all Wood Anomaly values given the incidence angles that are considered (Remark 2.1.5).

We use the solver of [13] for these computations, which is one of the periodic Green's function-based methods mentioned earlier (Section 2.2.1). This method calculates the amplitudes a_n of the Floquet series expansion

$$\begin{aligned} \mu(\boldsymbol{r})\sqrt{1+[f'(x)]^2} &\equiv \sum_{n=-\infty}^{\infty} a_n e^{i\alpha_n x} \\ &= \sum_{n=-\infty}^{\infty} a_n e^{i\left[k\sin(\theta)+n\frac{2\pi}{L}\right]x} \end{aligned} \tag{2.131}$$

(in this formula, θ is measured according to our convention as stated in Section 2.1.1 rather

than the convention of [13]). Using these amplitudes, we compute the Fourier amplitudes of $\mu_1(x)$ and $\mu_2(x)$ along with the real and imaginary parts of these functions for each case, and we plot the results.

Remark 2.3.5. *By the definition of $\mu_1(x)$ given in (2.88), the Fourier amplitudes of $\mu_1(x)$ are equal to $-\frac{a_n}{2ki}$. In agreement with this, the plots of the FFTs of $\mu_1(x)$ and $\mu_2(x)$ are appropriately scaled by the numbers of discretization points used to represent these functions.*

We note that the solver was implemented in FORTRAN 77 using "double precision" and "double complex" data types. Thus, the accuracy of its floating point arithmetic is approximately 16 digits; we will refer to this level of accuracy as "double precision accuracy" or "machine precision accuracy." As a result, only a subset of the Floquet modes can contribute to its numerical representation of $\mu(\boldsymbol{r})\sqrt{1+[f'(x)]^2}$. We call these the "significant" modes. Another result is that its calculation of the amplitudes a_n of modes which are not significant is entirely dominated by round-off error. This error carries over into the calculations of $\mu_1(x)$ and $\mu_2(x)$, and it can be seen in the plots of their Fourier amplitudes, where the insignificant modes have calculated amplitudes which are approximately 10^{-16} (with slight variation from case to case) in magnitude.

For each case where $k = 1.5 \times 2\pi$ (Cases 1, 3 and 5), $\mu_1(x)$ and $\mu_2(x)$ have similar Fourier spectra (Figures 2.11, 2.13 and 2.15). For $k = 100.5 \times 2\pi$, however, $\mu_1(x)$ and $\mu_2(x)$ differ strongly: in Case 2, in which there are only simple reflections, $\mu_2(x)$ oscillates much less than $\mu_1(x)$ does (Figure 2.12), while in Case 4 (with multiple reflections) and Case 6 (with multiple reflections and shadowing) $\mu_1(x)$ has many fewer significant Fourier modes than $\mu_2(x)$ has (Figures 2.14 and 2.16).

We also note that $\mu_1(x)$ has about three times as many significant modes in Case 2 than it has in Case 1, while the number of such modes for $\mu_2(x)$ is the same in both cases even though k is significantly larger in Case 2. This behavior is in agreement with the high-frequency ansatz described in Section 2.3.2.1 (the ansatz introduced in [17]) which motivated the formulation of $\mu_2(x)$ in Section 2.3.2.2. On the other hand, both $\mu_1(x)$ and $\mu_2(x)$ become increasingly oscillatory as k increases in the other cases.

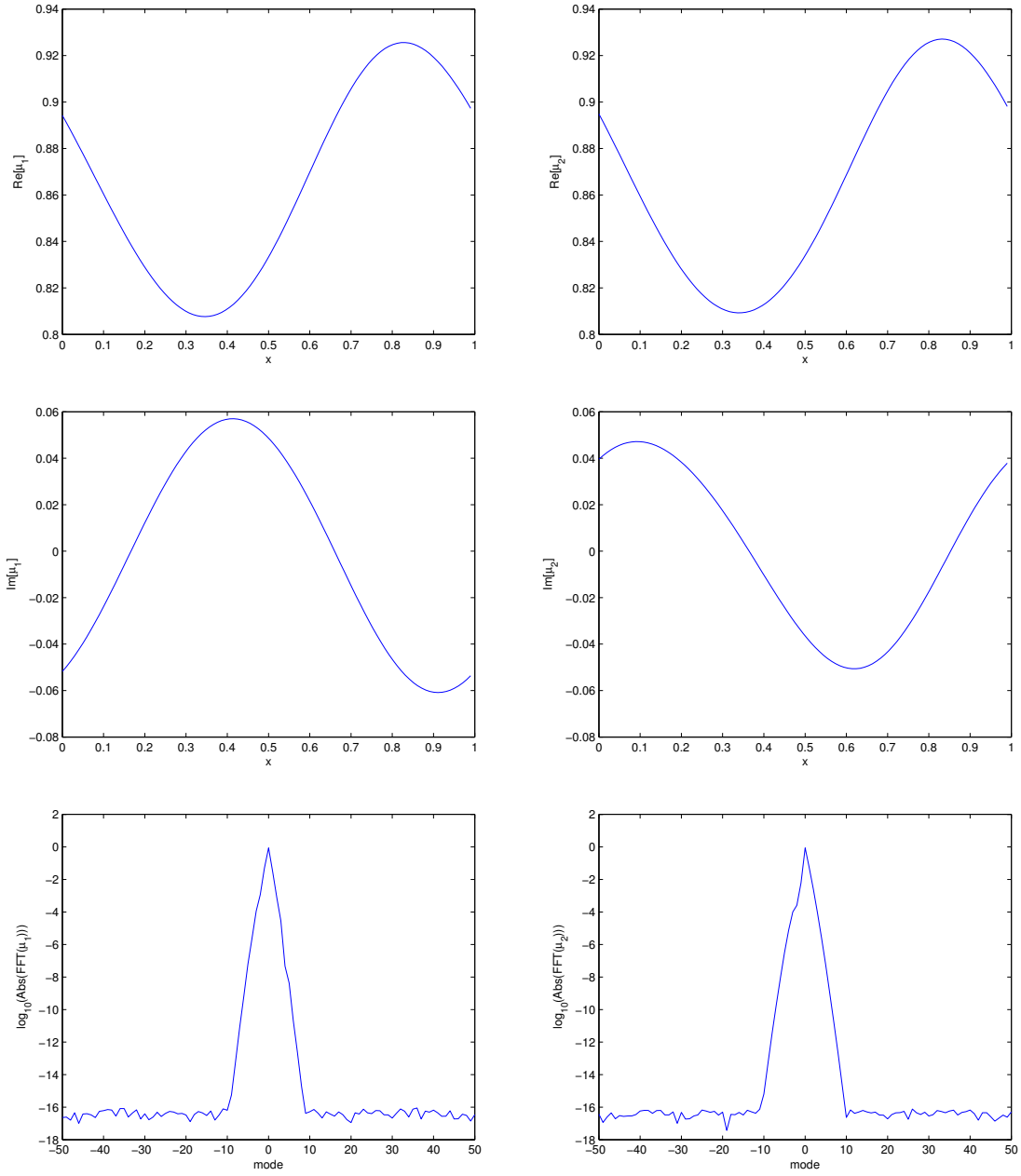


Figure 2.11: Case 1: real parts, imaginary parts and Fourier amplitudes of $\mu_1(x)$ and $\mu_2(x)$

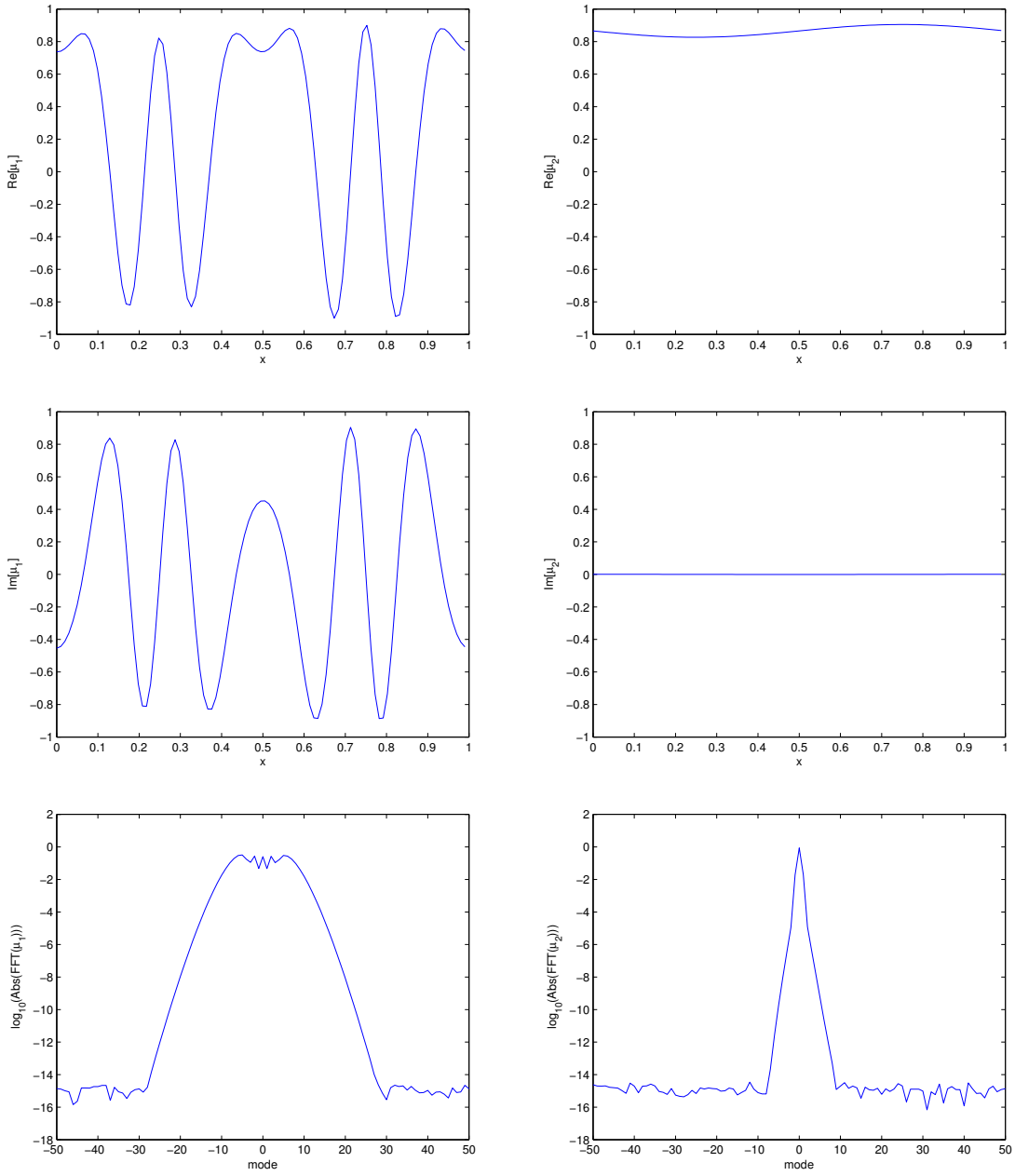


Figure 2.12: Case 2: real parts, imaginary parts and Fourier amplitudes of $\mu_1(x)$ and $\mu_2(x)$

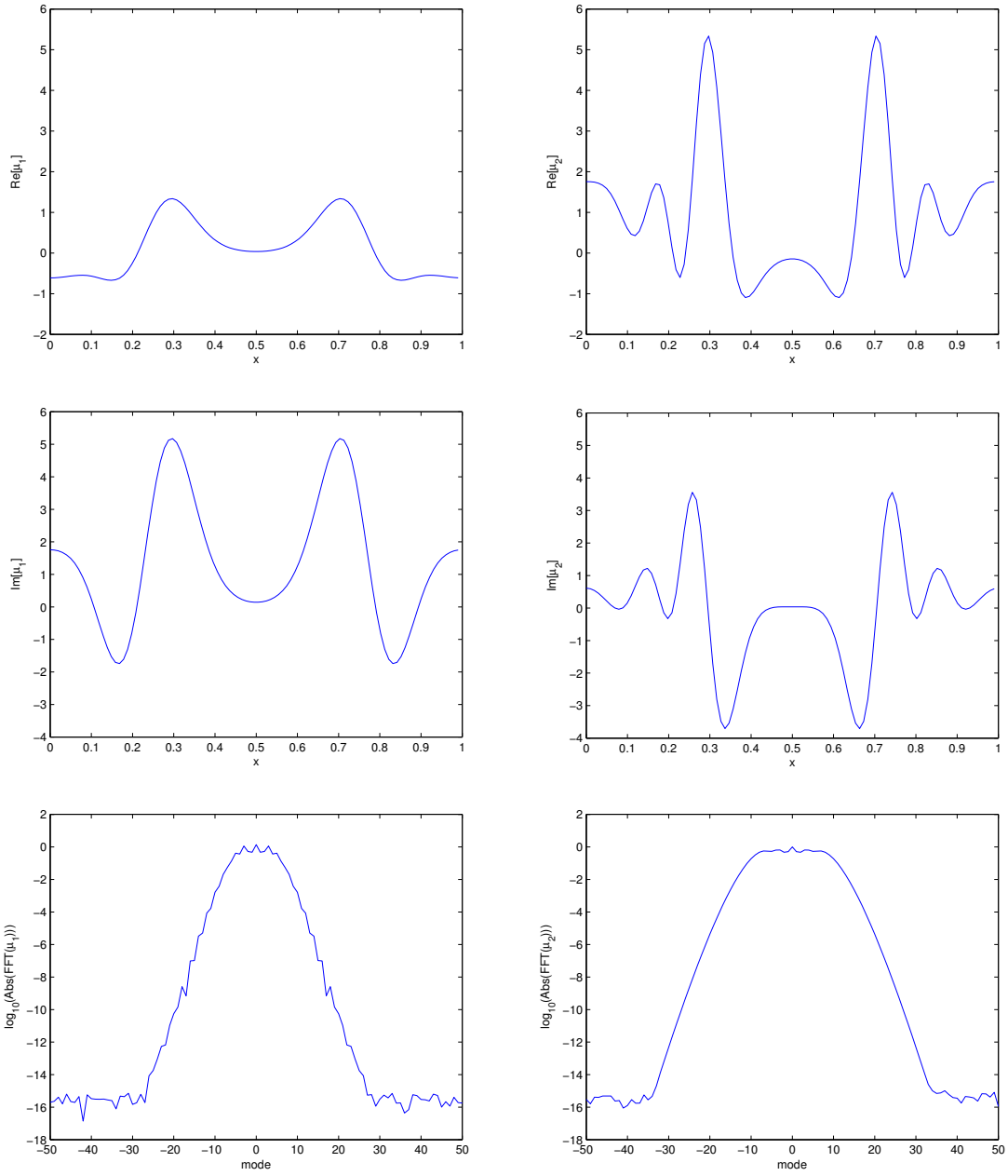


Figure 2.13: Case 3: real parts, imaginary parts and Fourier amplitudes of $\mu_1(x)$ and $\mu_2(x)$

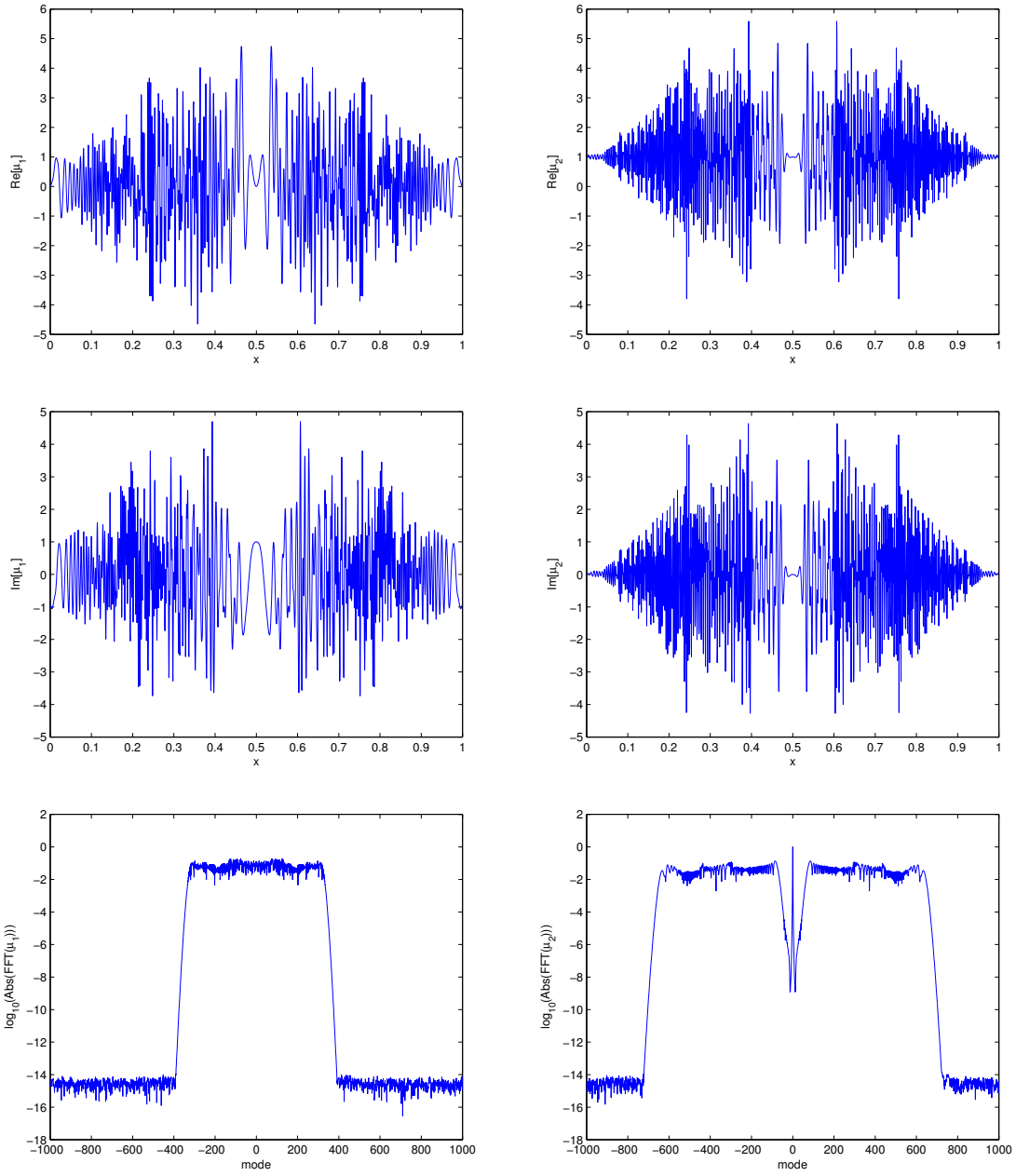


Figure 2.14: Case 4: real parts, imaginary parts and Fourier amplitudes of $\mu_1(x)$ and $\mu_2(x)$

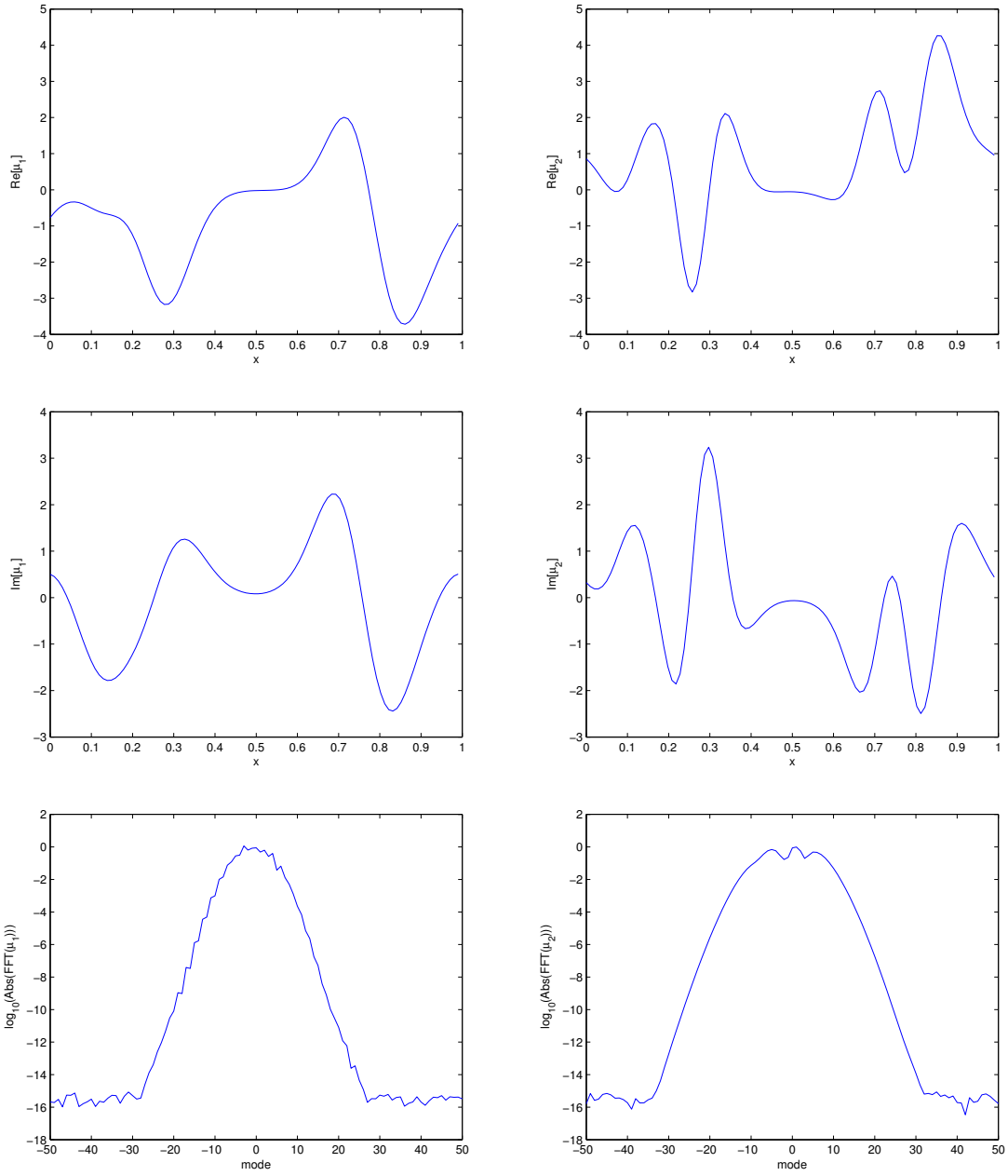


Figure 2.15: Case 5: real parts, imaginary parts and Fourier amplitudes of $\mu_1(x)$ and $\mu_2(x)$

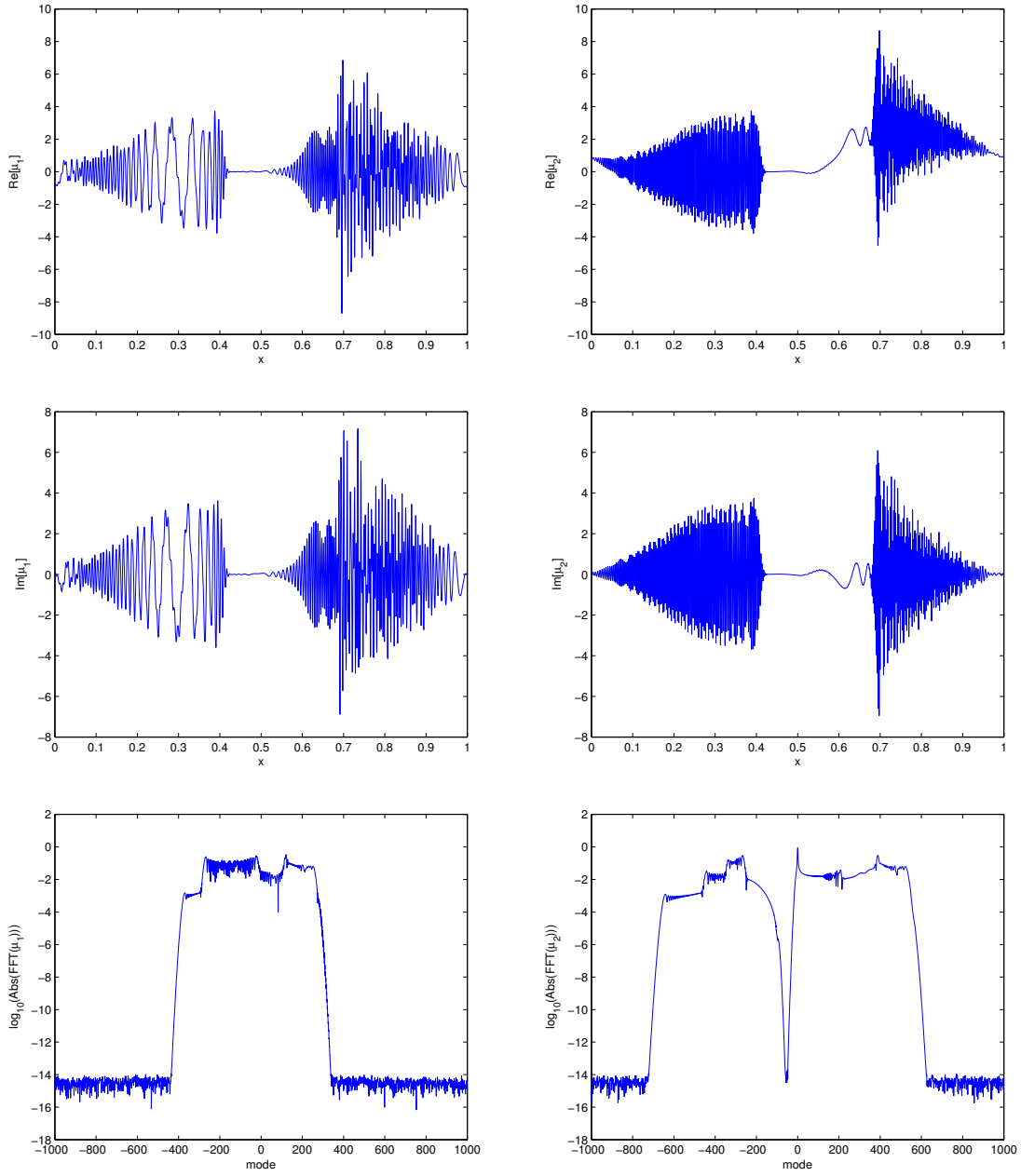


Figure 2.16: Case 6: real parts, imaginary parts and Fourier amplitudes of $\mu_1(x)$ and $\mu_2(x)$

2.3.3.6 Choice of Representations

As illustrated in the previous section's examples, it can be computationally advantageous to determine whether $\mu_1(x)$ or $\mu_2(x)$ will oscillate less for a given scattering configuration, since the lesser oscillating function requires fewer Fourier modes to be computed in order for accurate results to be achieved. In multiple-reflection cases, $\mu_1(x)$ required the same or fewer modes for its representation than $\mu_2(x)$ did. On the other hand, in simple-reflection cases, $\mu_2(x)$ had the same or fewer oscillations than $\mu_1(x)$ did, and it did not become increasingly oscillatory as k increased. Thus, motivated by these examples as well as our physical intuition, we choose to compute $\mu_2(x)$—i.e., solve (2.100)—if there are only simple reflections (as determined by the tests described earlier), while we compute $\mu_1(x)$—i.e., solve (2.90)—otherwise.

Remark 2.3.6. *Additional representations for the density $\mu(\boldsymbol{r})$ beyond the one containing $\mu_1(x)$ can be derived for configurations that give rise to multiple reflections. These expressions could be developed using the same kind of physical reasoning underlying the formulation of $\mu_2(x)$ for simple-reflection cases, so that then the integral equations of scattering could be solved for one or more slowly oscillating unknown functions even when multiple reflections (and perhaps also shadowing) occur. Such representations are not pursued in this thesis for the periodic rough surface problem. Appendix A, however, discusses how this reasoning is applied in the creation of an ansatz for the convex bounded obstacle problem (in which simple reflections and shadowing occur; see Sections A.2.1 and A.3.3 for details) as well as for the non-convex bounded obstacle problem (in which multiple scattering is also present; see Section A.6 for details).*

We note that requiring the computation of fewer Fourier modes of an L-periodic function is equivalent to requiring the computation of that function at fewer discretization points over the interval $x \in [0, L]$. The numerical method of this thesis uses discretizations in x; see Section 2.5 for further details.

2.4 Combining Windowing and the Two Representations

The algorithm proposed in this thesis results from the use of the smooth windowing functions and partitions of unity described in Section 2.2 combined with the oscillation-capturing

representations of the unknown densities $\mu(\boldsymbol{r})$ introduced in Section 2.3. In Section 2.3, the TE/sound-soft scattering equation (2.51) and the TM/sound-hard equation (2.52) (see also (2.35) and (2.39)) were each reformulated in two different ways, with the result being two pairs of equations: the pair of equations (2.90) for $\mu_1(x)$ (for problems in which multiple reflections arise) and the pair of equations (2.100) for $\mu_2(x)$ (for problems in which only simple reflections occur). Employing the approximation (2.57) of Section 2.2, then, gives rise to the approximating integral equations

$$\mu_m^A(x) \pm \int_{x-A}^{x+A} P_1(x,x',c,A) g(x,x') \frac{h(ku(x,x'))}{e^{iku(x,x')}} e^{ik\phi_m(x,x')} \mu_m^A(x')\, dx' = q_m(x), \quad 0 \le x \le L \tag{2.132}$$

($m = 1$ or $m = 2$). Here $\mu_m^A(x)$ are approximations of $\mu_m(x)$ that the method of this thesis further numerically approximates and computes.

In Section 3.1, we prove rigorously that the approximating integral operators inherent in (2.132) converge to the corresponding operators in (2.90) and (2.100) as $A \to \infty$ and that the solutions $\mu_m^A(x)$ therefore converge to $\mu_m(x)$ as $A \to \infty$. In particular, the convergence of the operators and the solutions is super-algebraic in A as $A \to \infty$ for problems with scattering surfaces $f(x)$ belonging to $C^\infty_{\text{per}}(L)$; this is a key property of our algorithm—one that contributes to its excellent performance.

2.5 Numerical Method

We present a numerical method for the computation of the integral equations (2.132). A numerical approximation to the solution of a given equation is calculated on an equispaced discretization of the interval $x \in [0, L]$ using a modified version of the Nyström method [37, 40]; see also [20], whose presentation we will refer to in Section 2.5.1. These modifications include the following: a splitting of the integral operator into four integrals in order to properly treat the integration over what is potentially more than one period of the scattering surface and the use of a potentially finer grid for the quadrature points than the one for the solution itself (Section 2.5.2). This leads to certain linear systems of equations that are solved numerically (Section 2.5.3). A special approach for the computation of the quadrature weights using a Chebyshev expansion is discussed in Section 2.5.3.1, and in Section 2.5.3.2 the implementation of our method via Fast Fourier Transforms and GMRES [53]

is developed. These techniques are especially useful for high-frequency cases; for simple-reflection cases, in particular, the computational time needed to solve the linear system is $\mathcal{O}(1)$ in k (Section 2.5.3.3). Finally, we briefly describe in Section 2.5.4 how the scattering efficiencies are computed. Thus, we obtain a spectrally accurate solver, which we will show in Section 3.2 to be convergent in the number of discretization points (super-algebraically convergent for configurations with smooth grating profiles) and will demonstrate in Chapter 4 to be highly efficient, accurate and fast for a wide variety of scattering configurations.

2.5.1 Spectrally Accurate Quadratures on Analytic Closed Curves

A Nyström method for the numerical solution of a boundary integral equation for the two-dimensional bounded obstacle scattering problem is presented in [20, pp. 67ff]. The configurations under consideration in that discussion contain obstacles that are impenetrable and have boundaries that are described either by analytic closed curves or by piecewise analytic closed curves with corners. For a case with a scattering surface of the first type, the analytic boundary curve of integration is parameterized in t with $0 \leq t \leq 2\pi$, and the integral equation to be solved is written in the form

$$\psi(t) - \int_0^{2\pi} K(t,\tau)\psi(\tau)\,d\tau = g(t), \quad 0 \leq t \leq 2\pi, \tag{2.133}$$

where (following [37, 40]; see also [35]) $K(t,\tau)$ is split as

$$K(t,\tau) = K_1(t,\tau)\log\left[4\sin^2\left(\frac{t-\tau}{2}\right)\right] + K_2(t,\tau) \tag{2.134}$$

with $K_1(t,\tau)$, $K_2(t,\tau)$ and $g(t)$ analytic. Given the $2n$ discretization points

$$t_j \equiv \frac{\pi j}{n}, \quad j = 0, 1, \ldots, 2n-1, \tag{2.135}$$

two quadrature formulas are used to approximate the integral. One is the trapezoidal rule

$$\int_0^{2\pi} f(\tau)\,d\tau \approx \frac{\pi}{n}\sum_{j=0}^{2n-1} f(t_j), \tag{2.136}$$

where here $f \in C^0_{\text{per}}(2\pi)$ is a generic integrand. The second,

$$\int_0^{2\pi} \log\left[4\sin^2\left(\frac{t-\tau}{2}\right)\right] f(\tau)\, d\tau \approx \sum_{j=0}^{2n-1} R_j^{(n)}(t) f(t_j), \quad 0 \le t \le 2\pi, \tag{2.137}$$

where

$$R_j^{(n)}(t) \equiv -\frac{2\pi}{n}\sum_{m=1}^{n-1}\frac{1}{m}\cos\left[m(t-t_j)\right] - \frac{\pi}{n^2}\cos\left[n(t-t_j)\right], \quad j = 0, 1, \ldots, 2n-1 \tag{2.138}$$

are the quadrature weights (as a function of t), handles the logarithmic singularity of the kernel at $\tau = t$. These formulas, which are obtained by substituting the trigonometric interpolation polynomial for f and integrating exactly, are spectrally accurate. Applying them, the unknown function $\psi(t)$ of (2.133) is approximated by the solution $\psi^{(n)}(t)$ of

$$\psi^{(n)}(t) - \sum_{j=0}^{2n-1}\left\{R_j^{(n)}(t)K_1(t,t_j) + \frac{\pi}{n}K_2(t,t_j)\right\}\psi^{(n)}(t_j) = g(t), \quad 0 \le t \le 2\pi, \tag{2.139}$$

and $\psi^{(n)}(t)$ is computed at the discretization points by solving the linear system

$$\psi^{(n)}(t_i) - \sum_{j=0}^{2n-1}\left\{R_j^{(n)}(t_i)K_1(t_i,t_j) + \frac{\pi}{n}K_2(t_i,t_j)\right\}\psi^{(n)}(t_j) = g(t_i), \quad i = 0, 1, \ldots, 2n-1. \tag{2.140}$$

Remark 2.5.1. *The quadrature formulas* (2.136) *and* (2.137) *converge exponentially for* 2π*-periodic* $f(t)$ *that are analytic, and it follows that the approximate solutions* $\psi^{(n)}(t)$ *converge uniformly and exponentially to* $\psi(t)$ *[20].*

2.5.2 Numerical Handling of the Integral Operators

The approach of the previous section cannot be directly applied to the integral operators found in the approximating equations (2.132). The support of the integrands (for each value of x, the support is equal to $2A$ due to the windowing function $P_1(x, x', c, A)$) may be larger than $2L$ (L is the period of both the scattering surface and the solution in each equation), and when this occurs a splitting of the kernels in a manner similar to the formula (2.134) would introduce artificial singularities at $x' = x - L$, $x' = x + L$, etc.

Instead, as we describe in Section 2.5.2.1, we partition the smooth windowing function

$P_1(x, x', c, A)$ using two smooth functions that are even about $x' = x$—one of which decays to zero within one period L of $x' = x$. Thus, we isolate the logarithmic singularity of the kernels at $x' = x$ in a manner that allows us to apply the quadrature formulas of the previous section to the resulting integral operators.

We develop the resulting spectrally accurate quadrature rule in detail in Section 2.5.2.2. Unlike the method described in the previous section, we allow for a discretization for the quadrature that is finer than the one used for the unknown functions $\mu_m^A(x)$ that are numerically approximated and computed—a property that we discuss further in Section 2.5.3.

2.5.2.1 Partitioning of the Integral Operators

The integral operators in the approximating equations (2.132) can be expressed as

$$\frac{L}{2\pi}\int_{t-\frac{2\pi A}{L}}^{t+\frac{2\pi A}{L}} K^m(t,\tau)\bar{\mu}_m^A(\tau)\,d\tau, \quad 0 \le t \le 2\pi$$

($m = 1$ or $m = 2$) where, using the parameterization

$$x(t) \equiv \frac{L}{2\pi}t, \tag{2.141}$$

we let

$$K^m(t,\tau) \equiv P_1\left[x(t), x(\tau), c, A\right] g\left[x(t), x(\tau)\right] \frac{h(ku\left[x(t), x(\tau)\right])}{e^{iku[x(t),x(\tau)]}} e^{ik\phi_m[x(t),x(\tau)]} \tag{2.142}$$

and

$$\bar{\mu}_m^A(t) \equiv \mu_m^A\left[x(t)\right]. \tag{2.143}$$

We note that a formula for $K^m(t,\tau)$ of the form (2.134) with analytic $K_1(t,\tau)$ and $K_2(t,\tau)$ does not hold. One reason for this is that the windowing function $P_1(x, x', c, A)$—while smooth—is not analytic; there are essential singularities at $x' = x \pm cA,\ x \pm A$, as can be seen by the formulas (2.53) and (2.54). More significantly, however, $x' = x + 2\pi Ll, \quad l = \pm 1, \pm 2, \ldots$ may lie within the support of $P_1(x, x', c, A)$ (there are two such values if $L < A < 2L$, four if $2L < A < 3L$, etc.), and the corresponding values $\tau = t + 2\pi l, \quad l = \pm 1, \pm 2, \ldots$ are singularities of $\log\left[4\sin^2\left(\frac{t-\tau}{2}\right)\right]$ but are not singularities of $K^m(t,\tau)$—a difficulty that does not arise for the integral operator in (2.133).

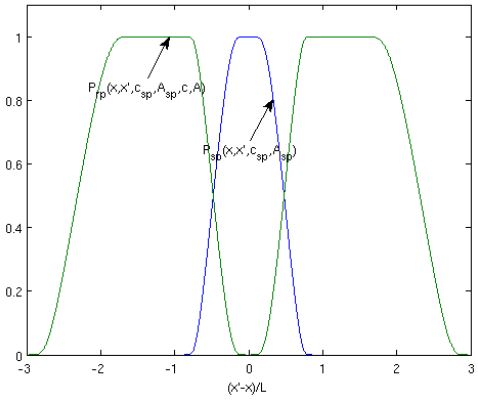


Figure 2.17: $P_{\rm sp}(x, x', c_{\rm sp}, A_{\rm sp})$ and $P_{\rm rp}(x, x', c_{\rm sp}, A_{\rm sp}, c, A)$, with $c_{\rm sp} = 0.01$, $A_{\rm sp} = \frac{7}{8}L$, $c = 0.5$ and $A = 3L$

In order to address this difficulty, we partition $P_1(x, x', c, A)$ into two windowing functions

$$P_1(x, x', c, A) = P_{\rm sp}(x, x', c_{\rm sp}, A_{\rm sp}) + P_{\rm rp}(x, x', c_{\rm sp}, A_{\rm sp}, c, A), \tag{2.144}$$

the first one of which can be used to isolate and treat the "singular part" of the kernel at $x' = x$; the remaining "regular part" of the kernel is multiplied by $P_{\rm rp}$. In particular, given (2.53) and (2.54), we let

$$P_{\rm sp}(x, x', c_{\rm sp}, A_{\rm sp}) \equiv S(x' - x, c_{\rm sp}A_{\rm sp}, A_{\rm sp}), \tag{2.145}$$

and

$$P_{\rm rp}(x, x', c_{\rm sp}, A_{\rm sp}, c, A) \equiv P_1(x, x', c, A) - P_{\rm sp}(x, x', c_{\rm sp}, A_{\rm sp}) \tag{2.146}$$

for certain values of $c_{\rm sp}$, $A_{\rm sp}$, c and A. These smooth windowing functions, centered about $x' = x$, are chosen using the constraints $A_{\rm sp} < L$ and $A_{\rm sp} < A$ so that $P_{\rm sp}(x, x', c_{\rm sp}, A_{\rm sp})$ vanishes for x' outside a subinterval of $(x - L, x + L)$ contained within the support of $P_1(x, x', c, A)$. Also, $c_{\rm sp}$ and c are chosen so that the partition functions do not decay very sharply (see Remark 2.2.2); after some numerical testing, we settled upon $c_{\rm sp} = 0.01$ and $c = 0.5$—the values used for all of the numerical results in Chapter 4—since these choices give satisfactory results. Figure 2.17 illustrates one example of this partition.

Using this partition, we write

$$K_{\mathrm{sp}}^{m}(t,\tau) \equiv P_{\mathrm{sp}}\left[x(t),x(\tau),c_{\mathrm{sp}},A_{\mathrm{sp}}\right] g\left[x(t),x(\tau)\right] \frac{h(ku\left[x(t),x(\tau)\right])}{e^{iku[x(t),x(\tau)]}} e^{ik\phi_m[x(t),x(\tau)]} \tag{2.147}$$

and

$$K_{\mathrm{rp}}^{m}(t,\tau) \equiv P_{\mathrm{rp}}\left[x(t),x(\tau),c_{\mathrm{sp}},A_{\mathrm{sp}},c,A\right] g\left[x(t),x(\tau)\right] \frac{h(ku\left[x(t),x(\tau)\right])}{e^{iku[x(t),x(\tau)]}} e^{ik\phi_m[x(t),x(\tau)]}. \tag{2.148}$$

Since the support of $K_{\mathrm{sp}}^{m}(t,\tau)$ lies within a subinterval of $(t-2\pi, t+2\pi)$, we may split this function according to the formula

$$K_{\mathrm{sp}}^{m}(t,\tau) = K_{\mathrm{sp1}}^{m}(t,\tau)\log\left[4\sin^2\left(\frac{t-\tau}{2}\right)\right] + K_{\mathrm{sp2}}^{m}(t,\tau), \tag{2.149}$$

where

$$\begin{aligned} &K_{\mathrm{sp1}}^{m}(t,\tau) \\ &\quad\equiv \frac{i}{\pi} P_{\mathrm{sp}}\left[x(t),x(\tau),c_{\mathrm{sp}},A_{\mathrm{sp}}\right] g\left[x(t),x(\tau)\right] ku\left[x(t),x(\tau)\right] \frac{J_1\left(ku\left[x(t),x(\tau)\right]\right)}{e^{iku[x(t),x(\tau)]}} e^{ik\phi_m[x(t),x(\tau)]} \\ &\quad= -\frac{k}{2\pi} P_{\mathrm{sp}}\left[x(t),x(\tau),c_{\mathrm{sp}},A_{\mathrm{sp}}\right] \frac{f\left[x(\tau)\right] - f\left[x(t)\right] - \left[x(\tau)-x(t)\right] f'\left[x(t)\right]}{u\left[x(t),x(\tau)\right]} \\ &\quad\times \frac{J_1\left(ku\left[x(t),x(\tau)\right]\right)}{e^{iku[x(t),x(\tau)]}} e^{ik\phi_m[x(t),x(\tau)]} \end{aligned} \tag{2.150}$$

($J_1(z)$ is the first-order Bessel function of the first kind) and

$$K_{\mathrm{sp2}}^{m}(t,\tau) \equiv K_{\mathrm{sp}}^{m}(t,\tau) - K_{\mathrm{sp1}}^{m}(t,\tau)\log\left[4\sin^2\left(\frac{t-\tau}{2}\right)\right]. \tag{2.151}$$

While $K_{\mathrm{sp1}}^{m}(t,\tau)$ and $K_{\mathrm{sp2}}^{m}(t,\tau)$ are not analytic, they are smooth for smooth profiles f. The same holds for $K_{\mathrm{rp}}^{m}(t,\tau)$; no splitting of this function is necessary.

Remark 2.5.2. *The differentiability properties of the functions $K_{sp1}^{m}(t,\tau)$, $K_{sp2}^{m}(t,\tau)$ and $K_{rp}^{m}(t,\tau)$ follow from the well-known expressions (see, for example, [3])*

$$J_1(x) = \sum_{p=0}^{\infty} \frac{(-1)^p}{p!(1+p)!}\left(\frac{x}{2}\right)^{1+2p} \tag{2.152}$$

and

$$Y_1(x) = \frac{2}{\pi}\left[\log\left(\frac{x}{2}\right) + C\right] J_1(x) - \frac{1}{\pi}\frac{2}{x} - \frac{1}{\pi}\sum_{p=0}^{\infty} \frac{(-1)^p}{p!(1+p)!}\left(\frac{x}{2}\right)^{1+2p} [\psi(p+1) + \psi(p)], \tag{2.153}$$

where here $\psi(0) \equiv 0$,

$$\psi(p) \equiv \sum_{m=1}^{p} \frac{1}{m}, \quad p = 1, 2, \ldots, \tag{2.154}$$

and

$$\gamma \equiv \lim_{p\to\infty}\left[\sum_{m=1}^{p} \frac{1}{m} - \log(p)\right], \tag{2.155}$$

which is Euler's constant. In particular,

$$K^m_{sp2}(t,t) = K^m_{sp}(t,t) = \frac{1}{2\pi}\frac{f''[x(t)]}{1 + \{f'[x(t)]\}^2}. \tag{2.156}$$

Thus, we express the integral operator of the approximating equations (2.132) as

$$\begin{aligned}
\frac{L}{2\pi}\int_{t-\frac{2\pi A}{L}}^{t+\frac{2\pi A}{L}} K^m(t,\tau)\bar{\mu}^A_m(\tau)\,d\tau &= \frac{L}{2\pi}\int_{t-\frac{2\pi A}{L}}^{t+\frac{2\pi A}{L}} K^m_{\mathrm{sp1}}(t,\tau)\log\left[4\sin^2\left(\frac{t-\tau}{2}\right)\right]\bar{\mu}^A_m(\tau)\,d\tau \\
&\quad + \frac{L}{2\pi}\int_{t-\frac{2\pi A}{L}}^{t+\frac{2\pi A}{L}} \left\{K^m_{\mathrm{sp2}}(t,\tau) + K^m_{\mathrm{rp}}(t,\tau)\right\}\bar{\mu}^A_m(\tau)\,d\tau \\
&= \frac{L}{2\pi}\int_{t-\frac{2\pi A_{\mathrm{sp}}}{L}}^{t+\frac{2\pi A_{\mathrm{sp}}}{L}} K^m_{\mathrm{sp1}}(t,\tau)\log\left[4\sin^2\left(\frac{t-\tau}{2}\right)\right]\bar{\mu}^A_m(\tau)\,d\tau \\
&\quad + \frac{L}{2\pi}\int_{t-\frac{2\pi A_{\mathrm{sp}}}{L}}^{t+\frac{2\pi A_{\mathrm{sp}}}{L}} K^m_{\mathrm{sp2}}(t,\tau)\bar{\mu}^A_m(\tau)\,d\tau \\
&\quad + \frac{L}{2\pi}\int_{t-\frac{2\pi A}{L}}^{t-\frac{2\pi c_{\mathrm{sp}} A_{\mathrm{sp}}}{L}} K^m_{\mathrm{rp}}(t,\tau)\bar{\mu}^A_m(\tau)\,d\tau \\
&\quad + \frac{L}{2\pi}\int_{t+\frac{2\pi c_{\mathrm{sp}} A_{\mathrm{sp}}}{L}}^{t+\frac{2\pi A}{L}} K^m_{\mathrm{rp}}(t,\tau)\bar{\mu}^A_m(\tau)\,d\tau, \quad 0 \le t \le 2\pi.
\end{aligned} \tag{2.157}$$

2.5.2.2 Quadrature Rule

Using the formulas of Section 2.5.1, we develop a spectrally accurate quadrature rule for the integral (2.157).

Let n_{i} be the number of discretization points per period of the scattering surface that are used for the quadrature (n_{i} is chosen to be even, just as $2n$ in Section 2.5.1 is even), and

call the points within the support of $K^m(t,\tau)$ (within the support of the smooth windowing function $P_1\left[x(t), x(\tau), c, A\right]$ for any given x) the "integration points"; we later will refer to n_{i} as the "number of integration points per period." Also,

$$n_{\mathrm{per}} \equiv \left\lceil \frac{A}{L} \right\rceil \tag{2.158}$$

is the number of whole periodic intervals of the grating to the right of $x = L$ which have at least some discretization points within $[L, L + A)$. Thus, we discretize a total of $2n_{\mathrm{per}} + 1$ intervals (corresponding to $x' \in [-n_{\mathrm{per}}L, (n_{\mathrm{per}} + 1)L]$) at a level of n_{i} points per period in order to compute the quadrature, i.e., we use

$$t_j \equiv \frac{2\pi j}{n_{\mathrm{i}}}, \quad j = -n_{\mathrm{per}} \cdot n_{\mathrm{i}}, -n_{\mathrm{per}} \cdot n_{\mathrm{i}} + 1, \ldots, (n_{\mathrm{per}} + 1)n_{\mathrm{i}} - 1. \tag{2.159}$$

We note that for any given t only a subset of these values (namely, those lying within the support of the windowing function $P_1\left[x(t), x(\tau), c, A\right]$) play a role in the discrete quadrature.

Due to the smooth decay to 0 of the kernel functions $K^m_{\mathrm{sp1}}(t,\tau)$, $K^m_{\mathrm{sp2}}(t,\tau)$ and $K^m_{\mathrm{rp}}(t,\tau)$ over certain intervals, they each can be periodically extended to a $2\pi n$-periodic function of τ for some positive integer n. These extensions can be represented by certain trigonometric interpolation polynomials, so the quadrature formulas (2.136) and (2.137)—appropriately modified for $2\pi n$-periodic continuous functions that are discretized using n_{i} points per 2π-interval—give us

$$\begin{aligned} \frac{L}{2\pi} \int_{t-\frac{2\pi A_{\mathrm{sp}}}{L}}^{t+\frac{2\pi A_{\mathrm{sp}}}{L}} K^m_{\mathrm{sp1}}(t,\tau) \log\left[4\sin^2\left(\frac{t-\tau}{2}\right)\right] \bar{\mu}^A_m(\tau)\, d\tau \\ \approx \frac{L}{2\pi} \sum_{\left\{j:\, |t-t_j| < \frac{2\pi A_{\mathrm{sp}}}{L}\right\}} R_j^{\left(\frac{n_{\mathrm{i}}}{2}\right)}(t) K^m_{\mathrm{sp1}}(t,t_j) \bar{\mu}^A_m(t_j), \end{aligned} \tag{2.160}$$

$$\frac{L}{2\pi} \int_{t-\frac{2\pi A_{\mathrm{sp}}}{L}}^{t+\frac{2\pi A_{\mathrm{sp}}}{L}} K^m_{\mathrm{sp2}}(t,\tau) \bar{\mu}^A_m(\tau)\, d\tau \approx \frac{L}{2\pi} \frac{2\pi}{n_{\mathrm{i}}} \sum_{\left\{j:\, |t-t_j| < \frac{2\pi A_{\mathrm{sp}}}{L}\right\}} K^m_{\mathrm{sp2}}(t,t_j) \bar{\mu}^A_m(t_j), \tag{2.161}$$

$$\frac{L}{2\pi} \int_{t-\frac{2\pi A}{L}}^{t-\frac{2\pi c_{\mathrm{sp}} A_{\mathrm{sp}}}{L}} K^m_{\mathrm{rp}}(t,\tau) \bar{\mu}^A_m(\tau)\, d\tau \approx \frac{L}{2\pi} \frac{2\pi}{n_{\mathrm{i}}} \sum_{\left\{j:\, t-\frac{2\pi A}{L} < t_j < t-\frac{2\pi c_{\mathrm{sp}} A_{\mathrm{sp}}}{L}\right\}} K^m_{\mathrm{rp}}(t,t_j) \bar{\mu}^A_m(t_j) \tag{2.162}$$

and

$$\frac{L}{2\pi}\int_{t+\frac{2\pi c_{\rm sp}A_{\rm sp}}{L}}^{t+\frac{2\pi A}{L}} K_{\rm rp}^m(t,\tau)\bar{\mu}_m^A(\tau)\,d\tau \approx \frac{L}{2\pi}\frac{2\pi}{n_{\rm i}} \sum_{\left\{j:\ t+\frac{2\pi c_{\rm sp}A_{\rm sp}}{L}<t_j<t+\frac{2\pi A}{L}\right\}} K_{\rm rp}^m(t,t_j)\bar{\mu}_m^A(t_j). \quad (2.163)$$

Therefore, we obtain the spectrally accurate quadrature rule

$$\begin{aligned}
&\frac{L}{2\pi}\int_{t-\frac{2\pi A}{L}}^{t+\frac{2\pi A}{L}} K^m(t,\tau)\bar{\mu}_m^A(\tau)\,d\tau \\
&\quad\approx \frac{L}{2\pi} \sum_{\left\{j:\ |t-t_j|<\frac{2\pi A_{\rm sp}}{L}\right\}} \left\{R_j^{\left(\frac{n_{\rm i}}{2}\right)}(t)K_{\rm sp1}^m(t,t_j) + \frac{2\pi}{n_{\rm i}}K_{\rm sp2}^m(t,t_j)\right\}\bar{\mu}_m^A(t_j) \\
&\quad + \frac{L}{2\pi}\frac{2\pi}{n_{\rm i}} \sum_{\left\{j:\ \frac{2\pi c_{\rm sp}A_{\rm sp}}{L}<|t-t_j|<\frac{2\pi A}{L}\right\}} K_{\rm rp}^m(t,t_j)\bar{\mu}_m^A(t_j), \quad 0\le t\le 2\pi
\end{aligned} \quad (2.164)$$

for the integral operators in (2.132).

Given this quadrature rule, we approximate the integral equations (2.132) by equations of the form

$$\begin{aligned}
&\bar{\mu}_m^{A,n_{\rm i}}(t) \pm \frac{L}{2\pi} \sum_{\left\{j:\ |t-t_j|<\frac{2\pi A_{\rm sp}}{L}\right\}} \left\{R_j^{\left(\frac{n_{\rm i}}{2}\right)}(t)K_{\rm sp1}^m(t,t_j) + \frac{2\pi}{n_{\rm i}}K_{\rm sp2}^m(t,t_j)\right\}\bar{\mu}_m^{A,n_{\rm i}}(t_j) \\
&\quad \pm \frac{L}{2\pi}\frac{2\pi}{n_{\rm i}} \sum_{\left\{j:\ \frac{2\pi c_{\rm sp}A_{\rm sp}}{L}<|t-t_j|<\frac{2\pi A}{L}\right\}} K_{\rm rp}^m(t,t_j)\bar{\mu}_m^{A,n_{\rm i}}(t_j) = q_m\left[x(t)\right], \quad 0\le t\le 2\pi,
\end{aligned} \quad (2.165)$$

where $\bar{\mu}_m^{A,n_{\rm i}}(t)$ are the approximations of $\bar{\mu}_m^A(t)$ that arise from using the quadrature rule with $n_{\rm i}$ integration points per period. We fully discretize these equations in order to generate linear systems that are solved numerically; as shown in the following section, it is possible (and, for efficiency, sometimes convenient) to approximate $\bar{\mu}_m^A(t)$ on coarser discretizations than the ones used for the quadrature.

2.5.3 Linear Systems of Equations

As mentioned in Section 2.5.1, an important component of the original Nyström method [20, 37, 40] is the use of a single set of points for both the discretization of the unknown density and the discrete quadrature. We have found, however, that it is often convenient in our

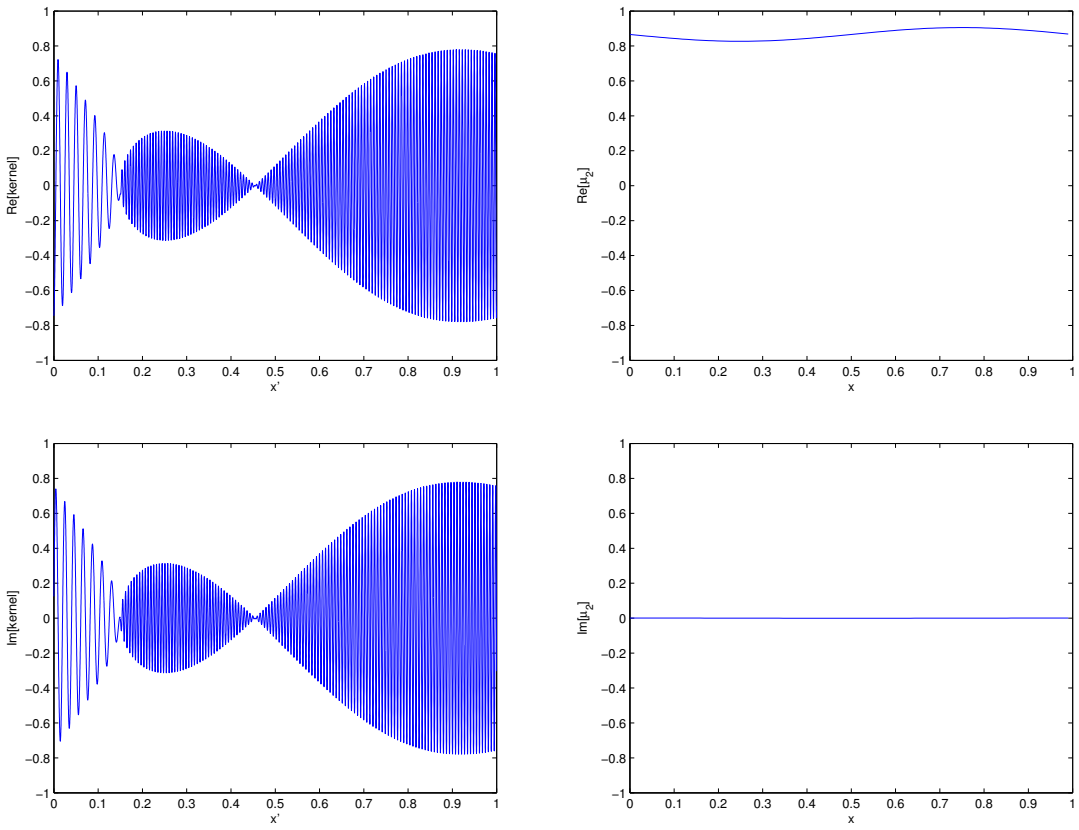


Figure 2.18: Real and imaginary parts of the kernel $g(x,x')\frac{h(ku(x,x'))}{e^{iku(x,x')}}e^{ik\phi_2(x,x')}$ $(x = 0.15)$ and density $\mu_2(x)$ for Case 2 of Section 2.3.3.5. These functions are from the TE/sound-soft integral equation of (2.100).

context to use fewer points for the unknown than for the quadrature. A striking example of this fact is given by the problem that arises by using the $\mu_2(x)$ solution in Case 2 of Section 2.3.3.5—a high-frequency simple-reflection case—which involves an integral kernel that is significantly more oscillatory than the solution; see Figure 2.18. As it happens, the same is true for many general configurations involving high or low frequencies and single or multiple scattering. Clearly, in all such cases it is advantageous to use smaller numbers of unknowns than quadrature points, since using a single set of points for both the quadrature and the discretization of the solution would be either inefficient or insufficient in each of these problems.

Thus, in our approach we approximate $\bar{\mu}_m^{A,n_\mathrm{i}}(t)$ at the n_t equispaced "target points" $t_{\ell\cdot n_\mathrm{mult}}$, $\ell = 0,1,2,\ldots,n_\mathrm{t}-1$ that form a subset of the set of quadrature points (2.159) which lie in $[0,2\pi)$, where n_t and the number n_i of integration points per period are even

integers and are related by the formula

$$n_{\rm i} = n_{\rm t} \times n_{\rm mult} \tag{2.166}$$

for an integer $n_{\rm mult} \geq 1$. Denoting $\tilde{\mu}_\ell$ as the approximate value of $\bar{\mu}_m^{A,n_{\rm i}}(t)$ at $t = t_{\ell \cdot n_{\rm mult}}$, we define $\hat{\mu}_j$ to be the values of the Fourier interpolation of $\tilde{\mu}_\ell,\ \ell = 0, 1, 2, \ldots, n_{\rm t} - 1$ onto the grid of quadrature points $t_j,\ j = 0, 1, 2, \ldots, n_{\rm i} - 1$ lying in $[0, 2\pi)$; these values of $\hat{\mu}_j$ are periodically extended to all $2n_{\rm per} + 1$ intervals used for the quadrature, i.e., to all of the points defined by (2.159). We approximate the finite sum of (2.165) at each target point $t = t_{\ell \cdot n_{\rm mult}}$ by the quantity

$$\begin{aligned}
&\sum_{j=0}^{n_{\rm i}-1} \hat{a}_{\ell,j} \hat{\mu}_j \\
&\quad \equiv \frac{L}{2\pi} \sum_{\left\{j:\ |\ell \cdot n_{\rm mult} - j| \leq n_{\rm hwsp}\right\}} \left\{ R^{\left(\frac{n_{\rm i}}{2}\right)}_{|\ell \cdot n_{\rm mult} - j|} K^m_{\rm sp1}(t_{\ell \cdot n_{\rm mult}}, t_j) + \frac{2\pi}{n_{\rm i}} K^m_{\rm sp2}(t_{\ell \cdot n_{\rm mult}}, t_j) \right\} \hat{\mu}_j \\
&\quad + \frac{L}{2\pi} \frac{2\pi}{n_{\rm i}} \sum_{\left\{j:\ n_{\rm hwspf}+1 \leq |\ell \cdot n_{\rm mult} - j| \leq n_{\rm hw}\right\}} K^m_{\rm rp}(t_{\ell \cdot n_{\rm mult}}, t_j) \hat{\mu}_j,
\end{aligned} \tag{2.167}$$

where

$$R^{\left(\frac{n_{\rm i}}{2}\right)}_j \equiv R^{\left(\frac{n_{\rm i}}{2}\right)}_j(0) = -\frac{4\pi}{n_{\rm i}} \sum_{m=1}^{\frac{n_{\rm i}}{2}-1} \frac{1}{m} \cos\left(\frac{m 2\pi j}{n_{\rm i}}\right) - \frac{(-1)^j 4\pi}{n_{\rm i}^2}, \quad j = 0, 1, \ldots, n_{\rm hwsp} \tag{2.168}$$

(following the notation of [20]) and

$$n_{\rm hw} \equiv \left\lfloor \frac{n_{\rm i} A}{L} \right\rfloor,\ n_{\rm hwsp} \equiv \left\lfloor \frac{n_{\rm i} A_{\rm sp}}{L} \right\rfloor,\ n_{\rm hwspf} \equiv \left\lfloor \frac{n_{\rm i} c_{\rm sp} A_{\rm sp}}{L} \right\rfloor; \tag{2.169}$$

we note that $\hat{a}_{\ell,j} = 0$ for certain values of ℓ, j if A is sufficiently small. Using this approximation, we therefore solve linear systems of the form

$$\tilde{\mu}_\ell \pm \sum_{j=0}^{n_{\rm i}-1} \hat{a}_{\ell,j} \hat{\mu}_j = q_m\left[x(t_{\ell \cdot n_{\rm mult}})\right], \quad \ell = 0, 1, 2, \ldots, n_{\rm t} - 1 \tag{2.170}$$

for $\tilde{\mu}_\ell,\ \ell = 0, 1, 2, \ldots, n_{\rm t} - 1$.

Remark 2.5.3. *Setting $n_i = n_t$ (i.e., $n_{mult} = 1$) results in the original Nyström approach*

for solving (2.132)*; for* $n_i = n_t$*, we have*

$$\tilde{\mu}_\ell = \hat{\mu}_\ell = \bar{\mu}_m^{A,n_i}(t_\ell), \quad \ell = 0, 1, 2, \ldots, n_t - 1. \tag{2.171}$$

If $n_{mult} > 1$*, the use of a Fourier interpolation of the values* $\tilde{\mu}_\ell$ *ensures spectral accuracy throughout our computations.*

We use Fast Fourier Transforms (FFTs) coupled with the iterative solver GMRES in order to calculate the unknowns $\tilde{\mu}_\ell$, $\ell = 0, 1, 2, \ldots, n_\mathrm{t} - 1$ of (2.170). The values $R^{\left(\frac{n_\mathrm{i}}{2}\right)}_{|\ell \cdot n_\mathrm{mult} - j|}$, $K^m_{\mathrm{sp1}}(t_{\ell \cdot n_\mathrm{mult}}, t_j)$, $K^m_{\mathrm{sp2}}(t_{\ell \cdot n_\mathrm{mult}}, t_j)$ and $K^m_{\mathrm{rp}}(t_{\ell \cdot n_\mathrm{mult}}, t_j)$ are computed once and are multiplied and added together according to (2.167) in order to form an array that is re-used at each GMRES iteration. In Section 2.5.3.1, an efficient method for computing the quadrature weights $R^{\left(\frac{n_\mathrm{i}}{2}\right)}_{|\ell \cdot n_\mathrm{mult} - j|}$ is developed, and the details regarding the FFTs and the use of GMRES are described in Section 2.5.3.2. Finally, we discuss in Section 2.5.3.3 how our approach gives rise to $\mathcal{O}(1)$ computational times for scattering problems in which only simple reflections occur.

2.5.3.1 Computing Quadrature Weights

In order to solve the systems (2.170), we compute the quadrature weights $R_j^{\left(\frac{n_\mathrm{i}}{2}\right)}$ (2.168) once and use them to construct an array for our iterative linear algebra solver. We note that each of these weights has $\mathcal{O}(n_\mathrm{i})$ terms, and n_i needs to increase as the wavenumber k increases in order to maintain accuracy in the numerical approximation.

One approach to handling the expense of calculating these weights is to precompute them for various values of j and n_i and store them in a table for future use, since they do not depend upon the physical parameters of the system. Such a table would have to be large in order to handle all of the cases of interest, however, especially if we wished to use it for high-frequency problems. A second approach—the one we take in this thesis—involves using Chebyshev expansions, and it is $\mathcal{O}(1)$ in n_i as $n_\mathrm{i} \to \infty$ (for fixed j).

Re-expressing (2.168) as

$$R_j^{\left(\frac{n_\mathrm{i}}{2}\right)} = -\frac{4\pi}{n_\mathrm{i}} \Re \left[\sum_{m=1}^{\frac{n_\mathrm{i}}{2}-1} \frac{1}{m} e^{i \frac{m 2\pi j}{n_\mathrm{i}}} \right] - \frac{(-1)^j 4\pi}{n_\mathrm{i}^2}, \tag{2.172}$$

we let

$$S_q(x) \equiv \sum_{m=1}^{\frac{n_\mathrm{i}}{2}-1} \frac{1}{m} e^{imx} \tag{2.173}$$

so that we may write the quadrature weights as

$$R_j^{\left(\frac{n_\mathrm{i}}{2}\right)} = -\frac{4\pi}{n_\mathrm{i}} \Re\left[S_q\left(\frac{2\pi j}{n_\mathrm{i}}\right)\right] - \frac{(-1)^j 4\pi}{n_\mathrm{i}^2}. \tag{2.174}$$

Using term-by-term differentiation and the formula for the sum of a geometric series, we have

$$\begin{aligned} S_q'(x) &= \sum_{m=1}^{\frac{n_\mathrm{i}}{2}-1} ie^{imx} \\ &= \begin{cases} ie^{ix} \frac{e^{i\left(\frac{n_\mathrm{i}}{2}-1\right)x}-1}{e^{ix}-1}, & x \neq 2\pi l,\ l = 0, \pm 1, \pm 2, \ldots \\ i\left(\frac{n_\mathrm{i}}{2}-1\right), & x = 2\pi l,\ l = 0, \pm 1, \pm 2, \ldots \end{cases} . \end{aligned} \tag{2.175}$$

So, we can reformulate $S_q(x)$ as an improper integral plus a constant:

$$\begin{aligned} S_q(x) &= \int_0^x S_q'(\xi)\, d\xi + S_q(0) \\ &= \int_0^x ie^{i\xi} \frac{e^{i\left(\frac{n_\mathrm{i}}{2}-1\right)\xi}-1}{e^{i\xi}-1}\, d\xi + \sum_{m=1}^{\frac{n_\mathrm{i}}{2}-1} \frac{1}{m}. \end{aligned} \tag{2.176}$$

Thus, we can use (2.176) to calculate $S_q\left(\frac{2\pi j}{n_\mathrm{i}}\right)$ and then (2.174) to compute $R_j^{\left(\frac{n_\mathrm{i}}{2}\right)}$.

Chebyshev expansions are employed in order to compute the integral in (2.176). A Fast Cosine Transform is used to efficiently calculate the coefficients of expansion of the function $S_q'(\xi)$ for $\xi \in \left[0, \frac{2\pi j}{n_\mathrm{i}}\right]$. For numerical stability, Taylor series approximations of the numerator and denominator of $S_q'(\xi)$ are computed for values of ξ near 0, with $i\left(\frac{n_\mathrm{i}}{2}-1\right)$ being used for $S_q'(0)$. Then, the Chebyshev coefficients of the integral in (2.176) are evaluated using a recurrence relationship for the coefficients of $S_q'(\xi)$. Since the range of integration $\left[0, \frac{2\pi j}{n_\mathrm{i}}\right]$ decreases as n_i increases (for fixed j), there is an upper bound on the number of zeros of $S_q'(x)$ in this interval. Thus, the number of Chebyshev coefficients required for a given accuracy of approximation of the integral in (2.176) is $\mathcal{O}(1)$ in n_i as $n_\mathrm{i} \to \infty$ (for fixed j). The last computational element for $S_q\left(\frac{2\pi j}{n_\mathrm{i}}\right)$, the sum $\sum_{m=1}^{\frac{n_\mathrm{i}}{2}-1} \frac{1}{m}$, diverges as $n_\mathrm{i} \to \infty$, and its direct evaluation requires $\mathcal{O}(n_\mathrm{i})$ computational time. However, given (2.155), for any

$\epsilon > 0$ there exists an integer $N > 0$ such that $\left|\sum_{m=1}^{\frac{n_\mathrm{i}}{2}-1} \frac{1}{m} - \log\left(\frac{n_\mathrm{i}}{2} - 1\right) - \gamma\right| < \epsilon$ if $n_\mathrm{i} > N$, where $\gamma \approx 0.5772$ is Euler's constant [3]. Therefore, the sum can be approximated to any desired degree of precision (e.g., up to machine precision) in $\mathcal{O}(1)$ steps as $n_\mathrm{i} \to \infty$; in practice, the cost of computing the sum directly for any n_i of interest is extremely small.

We note that for fixed A_sp the number $n_\mathrm{hwsp} + 1$ of quadrature weights to be computed increases as n_i increases. However, as indicated in Section 2.2.3 and as will be illustrated computationally in Section 4.3.1, as k (and thus n_i) increases for a given scattering profile $f(x)$ and incidence angle θ we may allow A (and thus A_sp) to decrease without sacrificing computational accuracy. Therefore, in practice $n_\mathrm{hwsp} + 1$ does not increase linearly with k. For high-frequency configurations in which only simple reflections arise, in particular, $n_\mathrm{hwsp} + 1$ has an upper bound as k increases; this will be shown in detail in Section 3.2.2.2.

2.5.3.2 FFT Interpolation and GMRES

Using the Chebyshev expansion-based method to calculate the quadrature weights, we compute the values $\hat{a}_{\ell,j}$ according to the formula (2.167)—recalling that $\hat{a}_{\ell,j} = 0$ for certain values of ℓ, j if A is sufficiently small. The left hand side of (2.170) can be re-formulated as

$$\tilde{\mu}_\ell \pm \sum_{j=0}^{n_\mathrm{i}-1} \hat{a}_{\ell,j}\hat{\mu}_j = \sum_{j=0}^{n_\mathrm{i}-1} a_{\ell,j}^{\pm}\hat{\mu}_j, \quad \ell = 0, 1, 2, \ldots, n_\mathrm{t} - 1, \tag{2.177}$$

where

$$a_{\ell,j}^{\pm} \equiv \begin{cases} 1 \pm \hat{a}_{\ell,j} & , \quad \ell \cdot n_\mathrm{mult} = j \\ \pm\hat{a}_{\ell,j} & , \quad \ell \cdot n_\mathrm{mult} \neq j \end{cases} \tag{2.178}$$

and (as stated previously) the values $\hat{\mu}_j,\ j = 0, 1, 2, \ldots, n_\mathrm{i} - 1$ are the Fourier interpolation of the values $\tilde{\mu}_\ell,\ \ell = 0, 1, 2, \ldots, n_\mathrm{t} - 1$ ($\hat{\mu}_{\ell \cdot n_\mathrm{mult}} = \tilde{\mu}_\ell$). We solve these linear systems for $\tilde{\mu}_\ell,\ \ell = 0, 1, 2, \ldots, n_\mathrm{t} - 1$ using the Krylov subspace solver GMRES [53]; for our calculations, we compute the elements $a_{\ell,j}^{\pm}$ once and re-use them at each GMRES iteration, and we employ an FFT-based method in order to perform the necessary Fourier interpolations (when $n_\mathrm{mult} > 1$).

The approach we take for the interpolations depends upon the distance between two adjacent target points $t_{\ell \cdot n_\mathrm{mult}}$ and $t_{(\ell+1) \cdot n_\mathrm{mult}}$ relative to the size (in periods) $\frac{A}{L}$ of the window of integration. If $t_{(\ell+1) \cdot n_\mathrm{mult}} - t_{\ell \cdot n_\mathrm{mult}} \leq \frac{2\pi A}{L}$—i.e., if the support of the window of integration (for a given target point) includes multiple target points—then all of the values $\hat{\mu}_j,\ j =$

$0, 1, 2, \ldots, n_{\mathrm{i}} - 1$ need to be computed in order to solve the linear systems. Thus, at each GMRES iteration, a forward FFT is applied to the n_{t}-element vector followed by appropriate zero padding and a reverse FFT in order to generate the n_{i}-element interpolation. However, if $t_{(\ell+1)\cdot n_{\mathrm{mult}}} - t_{\ell \cdot n_{\mathrm{mult}}} > \frac{2\pi A}{L}$—i.e., if the support of the window of integration only includes one target point—then certain $\hat{\mu}_j$ do not contribute to the matrix-vector product in the linear systems and therefore do not need to be computed; this holds, for example, for certain high-frequency problems in which A can be allowed to be very small, since if A is sufficiently small then $a^{\pm}_{\ell,j} = 0$ for certain j and all $\ell = 0, 1, 2, \ldots, n_{\mathrm{t}} - 1$. In such cases, a forward FFT of the n_{t}-element vector is performed at each GMRES iteration followed by inverse discrete Fourier transforms that are evaluated only for the relevant integration points t_j.

Remark 2.5.4. *The first and second of these conditions are equivalent to the conditions* $n_{mult} \leq n_{hw}$ *and* $n_{mult} > n_{hw}$*, respectively, by the formulas* (2.159)*,* (2.166) *and* (2.169)*.*

The linear systems are solved iteratively using unpreconditioned GMRES. At each step, approximations of the values $\tilde{\mu}_\ell$, $\ell = 0, 1, 2, \ldots, n_{\mathrm{t}} - 1$ are Fourier interpolated at the required integration points, the residual at the target points is computed, and new approximations of $\tilde{\mu}_\ell$ are calculated. The process is stopped when the residual is sufficiently small in norm; the threshold for this norm that is used for the results of Chapter 4 is 10^{-14}.

2.5.3.3 $\mathcal{O}(1)$ Computations for Simple-Reflection Cases

As discussed in Section 2.5.3.1, for a given scattering profile $f(x)$ and incidence angle θ we may allow A to decrease as k increases without sacrificing accuracy. In particular, for high-frequency simple-reflection cases—for which we compute the unknown $\mu_2(x)$ (2.101)—we may keep kA constant as k increases. In such cases, n_{hwsp} (2.169) does not need to increase even though k (and thus the number n_{i} of integration points per period) increases. In addition, for those problems we may fix the number n_{t} of target points used to approximate $\mu_2(x)$ as well as the number n_{hw} (2.169) of quadrature points that lie within half of the integration window about each target point. Given the computational procedures described in Section 2.5.3.2, it follows that our algorithm requires $\mathcal{O}(1)$ total computational time in order to solve the approximating linear systems (2.170) at a fixed accuracy level as $k \to \infty$ (for a given scattering profile $f(x)$ and incidence angle θ).

This property of our solver will be proved in detail in Section 3.2.2; computational results illustrating this fact will be provided in Section 4.3.1.1.

2.5.4 Computation of Scattering Efficiencies

Once the unknowns $\tilde{\mu}_\ell,\ \ell = 0, 1, 2, \ldots, n_\mathrm{t} - 1$ have been computed to a desired accuracy, they are Fourier interpolated onto the grid of integration points $t_j,\ j = 0, 1, 2, \ldots, n_\mathrm{i} - 1$. The coefficients B_n—given by either (2.42) (for TE/sound-soft problems) or (2.43) (for TM/sound-hard problems)—are then calculated using the trapezoidal rule (2.136) so that the scattering efficiencies e_n (2.44), in turn, can be generated. The sum of all of these efficiencies is also evaluated; given the "energy balance criterion" (2.45), the nearness of this sum to 1 serves as an indicator of the overall accuracy of our computations.

Chapter 3

Proofs

In Chapter 2, we developed a new algorithm to compute the scattering of plane waves by periodic, impenetrable rough surfaces. Our algorithm proceeds by

1. Recasting the TE/sound-soft and TM/sound-hard integral equations (2.35) and (2.39), which contain the periodic Green's function $\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')$ (2.21), as (2.51) and (2.52), which contain an infinite integral;

2. Using the finite integral approximation (2.57) that results from use of the smooth and gently-sloped windowing function $P_1(x, x', c, A)$ (2.54) supported in the interval $[x - A, x + A]$ (Section 2.2.2);

3. Representing the solution $\mu(\boldsymbol{r})$ in one of two different ways—for simple and multiple scattering configurations, respectively—which, as shown in Section 2.3.3.5, results in significant computational savings; and

4. Employing an efficient, spectrally accurate numerical method in order to discretize the approximating integral equations (2.132), as shown in Section 2.5.

In this chapter, we prove that the integral operators in these approximating equations are convergent as $A \to \infty$ (super-algebraically for smooth grating profiles) and thus that the corresponding solutions are convergent to the true solutions in the large A limit (Section 3.1). The main contributions in Section 3.1 are presented in Sections 3.1.1.1 and 3.1.3; certain necessary (but essentially well-known) existence and regularity results are discussed in this section as well. We also demonstrate that our numerical method is convergent (super-algebraically for smooth grating profiles) as the mesh-size tends to zero (Section 3.2.1). Note that, as shown in Chapter 4, often moderate and even small values of A and rather

coarse discretizations suffice to produce results of high quality. Finally, in Section 3.2.2 we show that the method requires $\mathcal{O}(1)$ operations (prescribed accuracies in fixed computational times for arbitrarily large k) for configurations that do not give rise to multiple reflections.

3.1 Properties of Analytical Approximation

The approximating integral equations (2.132) can be written in the form

$$(I \pm K_A^m)\,\mu_m^A(x) = q_m(x), \quad 0 \le x \le L. \tag{3.1}$$

Here m is either 1 (for multiple-reflection cases) or 2 (for simple-reflection cases), "+" is for TE/sound-soft scattering, "−" is for TM/sound-hard scattering, and

$$K_A^m \mu_m^A(x) \equiv \int_{x-A}^{x+A} P_1(x,x',c,A) g(x,x') \frac{h(ku(x,x'))}{e^{iku(x,x')}} e^{ik\phi_m(x,x')} \mu_m^A(x')\,dx'; \tag{3.2}$$

see (2.47), (2.53), (2.54), (2.85), (2.86), (2.87), (2.91), (2.102), (2.103) for the formulas of the various kernel functions. For L-periodic continuous functions $\varphi(x)$, let

$$K^m \varphi(x) \equiv \int_{-\infty}^{\infty} g(x,x') \frac{h(ku(x,x'))}{e^{iku(x,x')}} e^{ik\phi_m(x,x')} \varphi(x')\,dx' \tag{3.3}$$

for wavenumbers which are not Wood Anomaly values (Remark 2.1.5), i.e., for

$$k \neq \pm \left[k\sin(\theta) + n\frac{2\pi}{L} \right] \tag{3.4}$$

for any integer n. Then, we may write the exact scattering equations (2.90) and (2.100) as

$$(I \pm K^m)\,\mu_m(x) = q_m(x), \quad 0 \le x \le L, \tag{3.5}$$

and we have

$$(K^m - K_A^m)\,\varphi(x) = \int_{-\infty}^{\infty} P_2(x,x',c,A) g(x,x') \frac{h(ku(x,x'))}{e^{iku(x,x')}} e^{ik\phi_m(x,x')} \varphi(x')\,dx', \tag{3.6}$$

where $P_2(x,x',c,A)$ is given by (2.55).

In Section 3.1.1, we show that K_A^m converges in a certain norm to K^m as $A \to \infty$

for k which are not Wood Anomaly values, which implies that for such k the approximate solutions $\mu_m^A(x)$ exist for sufficiently large A and converge in that norm to $\mu_m(x)$ as $A \to \infty$. Also, we demonstrate that $\mu_m(x)$ and $\mu_m^A(x)$ belong to $C^\infty_{\text{per}}(L)$ for configurations with smooth grating surfaces (Section 3.1.2). Finally, we establish in Section 3.1.3 the super-algebraic convergence of the integral operators K_A^m (and thus of the approximate solutions) in such cases.

3.1.1 Convergence of Integral Operators and Solutions

We consider the Banach space of functions $\varphi \in C^0_{\text{per}}(L)$, i.e., $\varphi(x)$ is continuous and L-periodic (2.1), with norm

$$||\varphi||_\infty \equiv \max_{x\in[0,L]} |\varphi(x)|\,; \tag{3.7}$$

we note that convergence in this norm is equivalent to uniform convergence for $x \in [0, L]$. The approximating integral operator K_A^m can be re-expressed as the finite integral

$$K_A^m\varphi(x) = \int_0^L G_A^m(x, x', k, c)\varphi(x')\,dx', \quad 0 \le x \le L; \tag{3.8}$$

here

$$\begin{aligned} &G_A^m(x, x', k, c) \\ &\quad\equiv \sum_{n=-\infty}^{\infty} P_1(x, x' + nL, c, A)g(x, x' + nL)\frac{h(ku(x, x' + nL))}{e^{iku(x,x'+nL)}}e^{ik\phi_m(x,x'+nL)} \\ &\quad= \sum_{n=-N_A}^{N_A} P_1(x, x' + nL, c, A)g(x, x' + nL)\frac{h(ku(x, x' + nL))}{e^{iku(x,x'+nL)}}e^{ik\phi_m(x,x'+nL)}, \quad 0 \le x, x' \le L, \end{aligned} \tag{3.9}$$

where by (2.158)

$$N_A \equiv n_{\text{per}} = \left\lceil \frac{A}{L} \right\rceil. \tag{3.10}$$

Since the function $G_A^m(x, x', k, c)$ is "weakly singular," having a logarithmic singularity at $x' = x$ like the periodic Green's function on the grating surface (see Section 2.1.3.1 for details), K_A^m is a compact operator on $C^0_{\text{per}}(L)$ [35, Theorem 2.22].

We demonstrate in Section 3.1.1.1 that $||K^m - K_A^m||_\infty \to 0$ as $A \to \infty$ for k which are not Wood Anomaly values. Thus, for such k it holds that K^m also is a compact operator

on $C^0_{\text{per}}(L)$, and this leads to the result (as we prove in Section 3.1.1.2) that the solutions $\mu^A_m \in C^0_{\text{per}}(L)$ of the approximating equations (3.1) exist for sufficiently large A and converge in norm to the solutions $\mu_m \in C^0_{\text{per}}(L)$ of the exact equations (3.5) as $A \to \infty$ for such k.

3.1.1.1 Convergence of Integral Operators

In order to demonstrate the convergence in norm of K^m_A to K^m as $A \to \infty$, we begin by restating a key lemma from [16] which, for k which are not Wood Anomaly values, allows us to rewrite K^m as

$$K^m\varphi(x) = \int_0^L G^m(x, x', k)\varphi(x')\,dx', \quad 0 \le x \le L, \tag{3.11}$$

where

$$G^m(x, x', k) \equiv \sum_{n=-\infty}^{\infty} g(x, x' + nL)\frac{h(ku(x, x' + nL))}{e^{iku(x,x'+nL)}}e^{ik\phi_m(x,x'+nL)}, \quad 0 \le x, x' \le L. \tag{3.12}$$

Lemma 3.1.1. *Let s and μ be complex valued functions of bounded variation defined on the interval $[0, 1] \subset \mathbb{R}$ such that*

$$s(t), \mu(t) \to 0 \text{ as } t \to 0. \tag{3.13}$$

Write

$$\begin{aligned} s(t) &= a(t) + b(t), \\ \mu(t) &= c(t) + d(t), \end{aligned} \tag{3.14}$$

where the real and imaginary parts of a, b, c and d are monotone functions of t and

$$\begin{aligned} &\Re(a), \Im(a) \uparrow 0, \quad \Re(b), \Im(b) \downarrow 0, \\ &\Re(c), \Im(c) \uparrow 0, \quad \Re(d), \Im(d) \downarrow 0. \end{aligned} \tag{3.15}$$

Let

$$r_n = (nA - r) + s\left(\frac{1}{n}\right), \tag{3.16}$$

with $\Im(A) \ge 0$, $A \ne 2\pi l$ $(l \in \mathbb{Z})$, $r \in \mathbb{C}$ and s defined as above. Then,

1. *there exists a constant C such that*

$$\left|\sum_{n=1}^{N} e^{ir_n}\right| \leq C \text{ for all } N \geq 1; \tag{3.17}$$

2. *if $\mu_n = \mu\left(\frac{1}{n}\right)$ then the series*

$$\sum_{n=1}^{\infty} \mu_n e^{ir_n} \tag{3.18}$$

converges;

3. *if s and μ depend on an additional parameter $\lambda \in \mathbb{R}^p$ and the convergence in (3.15) is uniform for $\lambda \in I \subset \mathbb{R}^p$, then the convergence of the series (3.18) is uniform for $\lambda \in I$.*

With this lemma in view, let

$$\mathcal{H} \equiv \left\{(x, x', k) : -\frac{\gamma}{2} \leq x, x' \leq L + \frac{\gamma}{2},\ k \neq \pm\left[k\sin(\theta) + n\frac{2\pi}{L}\right] \text{ for } n \in \mathbb{Z}\right\} \tag{3.19}$$

for some small $\gamma > 0$, so that $-L - \gamma \leq x' - x \leq L + \gamma$ and k is by definition not a Wood Anomaly value for $(x, x', k) \in \mathcal{H}$. Then, the following theorem holds.

Theorem 3.1.1. *The series*

$$R^m \equiv \sum_{n=2}^{\infty} g(x, x' + nL)\frac{h(ku(x, x' + nL))}{e^{iku(x,x'+nL)}} e^{ik\phi_m(x,x'+nL)} \tag{3.20}$$

converges uniformly for (x, x', k) in compact subsets of $\mathcal{H}$.

Proof. By Taylor series expansions of the formulas (2.47) and (2.85) for $u(x, x')$ and $g(x, x')$, the formula $h(t) \equiv tH_1^1(t)$ of (2.86) and the asymptotic formula (2.62) for $H_1^1(z)$, we have

$$g(x, x' + nL)\frac{h(ku(x, x' + nL))}{e^{iku(x,x'+nL)}} \sim -\frac{i}{2}\sqrt{\frac{2k}{\pi}}e^{-i\frac{3\pi}{4}}\frac{f'(x)}{\sqrt{x' - x + nL}} + \mathcal{O}\left(\frac{1}{n^{\frac{3}{2}}}\right), \quad \text{as } n \to \infty. \tag{3.21}$$

Since a series containing just the $\mathcal{O}\left(\frac{1}{n^{\frac{3}{2}}}\right)$ terms converges absolutely and uniformly for (x, x', k) in compact subsets of $\mathcal{H}$, we focus on the $\mathcal{O}\left(\frac{1}{n^{\frac{1}{2}}}\right)$ terms and let

$$\mu(x, x', k, t) \equiv -\frac{i}{2}\sqrt{\frac{2k}{\pi}}e^{-i\frac{3\pi}{4}}\frac{f'(x)}{\sqrt{x' - x + \frac{L}{t}}}. \tag{3.22}$$

Also, (2.91) and (2.103) can be rewritten as

$$\phi_1(x,x'+nL) = [1+\sin(\theta)]\,nL + [1+\sin(\theta)]\,(x'-x) + \frac{[f(x')-f(x)]^2}{\sqrt{(x'-x+nL)^2+[f(x')-f(x)]^2}+(x'-x+nL)} \tag{3.23}$$

and

$$\phi_2(x,x'+nL) = [1+\sin(\theta)]\,nL + [1+\sin(\theta)]\,(x'-x) - \cos(\theta)\left[f(x')-f(x)\right] + \frac{[f(x')-f(x)]^2}{\sqrt{(x'-x+nL)^2+[f(x')-f(x)]^2}+(x'-x+nL)}. \tag{3.24}$$

Accordingly, we let

$$s(x,x',k,t) \equiv k\frac{[f(x')-f(x)]^2}{\sqrt{\left(x'-x+\frac{L}{t}\right)^2+[f(x')-f(x)]^2}+\left(x'-x+\frac{L}{t}\right)}. \tag{3.25}$$

Now, both $\mu(x,x',k,t)$ and $s(x,x',k,t)$ are well defined and continuously differentiable with respect to t for $t \in \left(0,\frac{1}{2}\right]$ (for sufficiently small γ), and they vanish as $t \to 0$. In particular, the improper integrals $\int_0^{\frac{1}{2}} \left|\frac{\partial \mu}{\partial t}\right| dt$ and $\int_0^{\frac{1}{2}} \left|\frac{\partial s}{\partial t}\right| dt$ exist (i.e., are finite), so $\mu(x,x',k,t)$ and $s(x,x',k,t)$ (defining $\mu(x,x',k,0)=0$ and $s(x,x',k,0)=0$) are functions of bounded variation for $t \in \left[0,\frac{1}{2}\right]$ [25]. Furthermore, both $\mu(x,x',k,t) \to 0$ and $s(x,x',k,t) \to 0$ uniformly for (x,x',k) in compact subsets of $\mathcal{H}$ as $t \to 0$. Finally, for

$$r_n \equiv k\phi_m(x,x'+nL) = (nA-r)+s\left(\frac{1}{n}\right) \tag{3.26}$$

we have $A = [1+\sin(\theta)]\,kL$, and for k which are not Wood Anomaly values the relation $A \neq 2\pi l$ $(l \in \mathbb{Z})$ holds. By Lemma 3.1.1, we conclude that

$$\sum_{n=2}^{\infty} \mu\left(x,x',k,\frac{1}{n}\right) e^{ir_n} \tag{3.27}$$

and therefore the series R^m converges uniformly for (x,x',k) in compact subsets of $\mathcal{H}$. $\square$

By a similar analysis, it can be shown that the series

$$\sum_{n=-\infty}^{-2} g(x,x'+nL)\frac{h(ku(x,x'+nL))}{e^{iku(x,x'+nL)}}e^{ik\phi_m(x,x'+nL)}$$

also converges uniformly for (x,x',k) in compact subsets of $\mathcal{H}$. Therefore, K^m indeed may be written as (3.11) for k which are not Wood Anomaly values.

Having this result, we next show that $|G^m(x,x',k)-G^m_A(x,x',k,c)|\to 0$ uniformly for (x,x',k) in compact subsets of $\mathcal{H}$ as $A\to\infty$. This is done by proving an extension of Lemma 3.1.1 and then applying this extension to the tails of the series $G^m_A(x,x',k,c)$ and $G^m(x,x',k)$.

Lemma 3.1.2. *Let $s(t)$, $\mu(t)$, μ_n and r_n be as in Lemma 3.1.1, including that s and μ depend on an additional parameter $\lambda\in\mathbb{R}^p$ with the convergence in (3.15) being uniform for $\lambda\in I\subset\mathbb{R}^p$. Let $w(A,t)$ be a real function that is monotone in t and defined for $(A,t)\in[A_0,\infty)\times[0,1]$ such that*

$$w(A,t)=1,\quad t\in[f(A),1] \tag{3.28}$$

and $w(A,t)\to 0$ as $t\to 0$, where $f(A)$ is a real monotone function such that $0<f(A)<1$ for $A\in[A_0,\infty)$ and $f(A)\downarrow 0$ as $A\to\infty$. Then,

1. *if $w^A_n=w\left(A,\frac{1}{n}\right)$, then the series*

$$\sum_{n=1}^{\infty} w^A_n\mu_n e^{ir_n} \tag{3.29}$$

   converges uniformly for $A\in[A_0,\infty)$;

2. *if w additionally depends on λ and $w(A,t)\to 0$ as $t\to 0$ uniformly for $\lambda\in I$, then the series (3.29) converges uniformly for $(\lambda,A)\in I\times[A_0,\infty)$, and it converges uniformly for $\lambda\in I$ to*

$$\sum_{n=1}^{\infty}\mu_n e^{ir_n} \tag{3.30}$$

   as $A\to\infty$.

Proof. We consider each part of the theorem in turn.

Part 1: Since $w(A,t)$ is a real monotone function of t on $[0,1]$ with $w \downarrow 0$ as $t \to 0$, by the conditions on $\mu(t)$ in Lemma 3.1.1 it follows that $w(A,t)\mu(t)$ is a complex valued function of bounded variation on $[0,1]$ with the property $w(A,t)\mu(t) \to 0$ as $t \to 0$ and can be written as

$$w(A,t)\mu(t) = w(A,t)c(t) + w(A,t)d(t), \tag{3.31}$$

where $w(A,t)c(t)$ and $w(A,t)d(t)$ are monotone functions of t such that

$$\Re\left[w(A,t)c(t)\right], \Im\left[w(A,t)c(t)\right] \uparrow 0, \quad \Re\left[w(A,t)d(t)\right], \Im\left[w(A,t)d(t)\right] \downarrow 0. \tag{3.32}$$

Thus, the series (3.29) converges. Since for each $t \in [0,1]$ the inequalities $0 \leq w(A,t) \leq 1$ hold for all $A \in [A_0, \infty)$, it follows from Part 3 of Lemma 3.1.1 that the convergence in (3.32) is uniform for $A \in [A_0, \infty)$, and thus the series (3.29) converges uniformly for $A \in [A_0, \infty)$.

Part 2: If w additionally depends on λ and $w(A,t) \to 0$ uniformly for $\lambda \in I$ as $t \to 0$, then it follows from Part 3 of Lemma 3.1.1 that the convergence in (3.32) is uniform for $\lambda \in I$ and thus that the series (3.29) converges uniformly for $(\lambda, A) \in I \times [A_0, \infty)$. Given the formula (3.28) for $w(A,t)$, it holds that for each $A \in [A_0, \infty)$ we have $w_n^A = 1$ for $n = 1, 2, 3, \ldots, M_A$ for

$$M_A \equiv \left\lfloor \frac{1}{f(A)} \right\rfloor. \tag{3.33}$$

In particular, M_A increases monotonically as A increases, and $M_A \to \infty$ as $A \to \infty$. Thus, by the uniform convergence properties of the two series (3.29) and (3.30), for every $\epsilon > 0$ and $\delta > 0$ there exists a real $A_{\epsilon,\delta}$ such that

$$\left| \sum_{n=1}^{\infty} \mu_n e^{ir_n} - \sum_{n=1}^{M_A} \mu_n e^{ir_n} \right| < \epsilon \tag{3.34}$$

and

$$\begin{aligned} \left| \sum_{n=1}^{\infty} w_n^A \mu_n e^{ir_n} - \sum_{n=1}^{M_A} \mu_n e^{ir_n} \right| &= \left| \sum_{n=1}^{\infty} w_n^A \mu_n e^{ir_n} - \sum_{n=1}^{M_A} w_n^A \mu_n e^{ir_n} \right| \\ &< \delta \end{aligned} \tag{3.35}$$

for $A > A_{\epsilon,\delta}$ and $\lambda \in I$. By the triangle inequality,

$$\begin{aligned}\left|\sum_{n=1}^{\infty}\mu_n e^{ir_n} - \sum_{n=1}^{\infty} w_n^A \mu_n e^{ir_n}\right| &\leq \left|\sum_{n=1}^{\infty}\mu_n e^{ir_n} - \sum_{n=1}^{M_A}\mu_n e^{ir_n}\right| \\ &\quad + \left|\sum_{n=1}^{M_A}\mu_n e^{ir_n} - \sum_{n=1}^{\infty} w_n^A \mu_n e^{ir_n}\right| \\ &< \epsilon + \delta.\end{aligned} \tag{3.36}$$

Therefore, the series (3.29) converges uniformly for $\lambda \in I$ to (3.30) as $A \to \infty$. □

Theorem 3.1.2. *For the series $G_A^m(x,x',k,c)$ and $G^m(x,x',k)$ given by (3.9) and (3.12), respectively, $|G^m(x,x',k) - G_A^m(x,x',k,c)| \to 0$ uniformly for (x,x',k) in compact subsets of $\mathcal{H}$ as $A \to \infty$.*

Proof. We recall that the series

$$R^m \equiv \sum_{n=2}^{\infty} g(x,x'+nL)\frac{h(ku(x,x'+nL))}{e^{iku(x,x'+nL)}}e^{ik\phi_m(x,x'+nL)} \tag{3.37}$$

converges uniformly for (x,x',k) in compact subsets of $\mathcal{H}$ by Theorem 3.1.1. Now, the windowing function $P_1\left(x,x'+\frac{L}{t},c,A\right)$—equal to 1 for $\left|x'-x+\frac{L}{t}\right| \leq cA$—satisfies the conditions on $w(A,t)$ in Lemma 3.1.2 for $t \in \left(0,\frac{1}{2}\right]$ and A sufficiently large. Thus, the series

$$\begin{aligned}R_A^m &\equiv \sum_{n=2}^{\infty} P_1(x,x'+nL,c,A)g(x,x'+nL)\frac{h(ku(x,x'+nL))}{e^{iku(x,x'+nL)}}e^{ik\phi_m(x,x'+nL)} \\ &= \sum_{n=2}^{N_A} P_1(x,x'+nL,c,A)g(x,x'+nL)\frac{h(ku(x,x'+nL))}{e^{iku(x,x'+nL)}}e^{ik\phi_m(x,x'+nL)}\end{aligned} \tag{3.38}$$

converges uniformly for such (x,x',k) to R^m as $A \to \infty$. Similarly, the series

$$\sum_{n=-\infty}^{-2} P_1(x,x'+nL,c,A)g(x,x'+nL)\frac{h(ku(x,x'+nL))}{e^{iku(x,x'+nL)}}e^{ik\phi_m(x,x'+nL)}$$

uniformly converges to

$$\sum_{n=-\infty}^{-2} g(x,x'+nL)\frac{h(ku(x,x'+nL))}{e^{iku(x,x'+nL)}}e^{ik\phi_m(x,x'+nL)}$$

as $A \to \infty$. Therefore, $|G^m(x,x',k) - G_A^m(x,x',k,c)| \to 0$ uniformly for (x,x',k) in compact subsets of $\mathcal{H}$ as $A \to \infty$. □

This leads to our main result of this section:

Theorem 3.1.3. *For k in compact subsets of the set of k which are not Wood Anomaly values, the operator K_A^m given by (3.8) and the operator K^m given by (3.11), we have $||K^m - K_A^m||_\infty \to 0$ as $A \to \infty$.*

Proof. By Theorem 3.1.2, for every $\epsilon > 0$ there exists a value A_ϵ such that

$$\begin{aligned} ||(K^m - K_A^m)\,\varphi||_\infty &= \max_{x\in[0,L]} \left| \int_0^L \left[G^m(x,x',k) - G_A^m(x,x',k,c)\right] \varphi(x')\,dx' \right| \\ &\leq \int_0^L \max_{x\in[0,L]} \left|G^m(x,x',k) - G_A^m(x,x',k,c)\right| \left|\varphi(x')\right|\,dx' \\ &< \epsilon \max_{x'\in[0,L]} \left|\varphi(x')\right| \int_0^L 1\,dx' \\ &= \epsilon L\,||\varphi||_\infty \end{aligned} \tag{3.39}$$

for $A > A_\epsilon$ and $\varphi \in C^0_{\text{per}}(L)$. Therefore,

$$||K^m - K_A^m||_\infty = \sup_{\varphi\in C^0_{\text{per}}(L)} \frac{||(K^m - K_A^m)\,\varphi||_\infty}{||\varphi||_\infty} \to 0 \tag{3.40}$$

as $A \to \infty$. □

Remark 3.1.1. *The convergence in norm of K_A^m to K^m as $A \to \infty$ holds for arbitrarily large k that are not Wood Anomaly values. We will demonstrate numerically in Section 4.3.1 that we may allow A to decrease as k increases (for a given scattering profile $f(x)$ and incidence angle θ) and maintain a desired level of computational accuracy in our results, as was illustrated in the simplified example of Section 2.2.3. Additionally, in Section 4.3.4 we will present a study of the application of our numerical method to problems with wavenumbers near and even at Wood Anomaly values, and we will show that our method is, in fact, effective for many of these cases; excellent results for yet more configurations with Wood Anomaly values of k—including a demonstration of the agreement between our computations of the scattering efficiencies and those made by other algorithms—are given in Appendices B and C.*

3.1.1.2 Existence and Convergence of Solutions

Since K_A^m is a compact operator on $C^0_{\text{per}}(L)$ (Section 3.1.1) and $||K^m - K_A^m||_\infty \to 0$ as $A \to \infty$ for k which are not Wood Anomaly values (Theorem 3.1.3), it follows that K^m also is a compact operator on $C^0_{\text{per}}(L)$ for such k [35, Theorem 2.17].

Because K^m (and thus $-K^m$) is compact, both the existence and the uniqueness of the solutions $\mu_m(x)$ of (3.5) can be established by showing that the corresponding homogeneous equations

$$(I \pm K^m)\,\varphi(x) = 0 \tag{3.41}$$

only have the trivial solution $\varphi(x) = 0$ (i.e., that $I \pm K^m$ are injective), since $I \pm K^m$ are injective if and only if they are surjective and since being injective implies that they have bounded inverses [35, Theorem 3.4]. This is a fundamental result of the theory developed in Riesz' [51] generalization of Fredholm's [27] work to Banach spaces; see, e.g., [35, Chapter 3] for further discussion of the Riesz theory. The injectivity of these operators, in turn, follows [43] from uniqueness of solutions of TE/sound-soft and TM/sound-hard problems for the Helmholtz equation above and below the scattering profile. The uniqueness of solutions of the latter problems has been established for many, but not all, configurations; see Remark 2.1.2. In any case, the injectivity of the integral operators above is assumed throughout this thesis.

We now prove the existence and convergence of our approximate solutions $\mu_m^A(x)$.

Theorem 3.1.4. *For k which are not Wood Anomaly values, let K_A^m be given by (3.8) and K^m be given by (3.11). Then, for such k, the solutions $\mu_m^A \in C^0_{per}(L)$ of the approximating scattering equations (3.1) exist for sufficiently large A, and these solutions converge uniformly on $[0, L]$ to the solutions $\mu_m \in C^0_{per}(L)$ of the exact scattering equations (3.5) as $A \to \infty$.*

Proof. Since K^m (and thus $-K^m$) are compact operators on $C^0_{\text{per}}(L)$, $I \pm K^m$ are bounded operators on $C^0_{\text{per}}(L)$. As stated previously, we assume that $I \pm K^m$ are injective and thus have a bounded inverse. Because K_A^m converges in norm to K^m as $A \to \infty$ (Theorem 3.1.3), it follows that $I \pm K_A^m$ have bounded inverses on $C^0_{\text{per}}(L)$ (i.e., the solutions $\mu_m^A \in C^0_{\text{per}}(L)$ exist) for sufficiently large A [35, Theorem 10.1]; by [35, Corollary 10.3], it also follows that

the error estimate

$$\left|\left|\mu_m - \mu_m^A\right|\right|_\infty \leq C \left|\left|\left[(I \pm K^m) - (I \pm K_A^m)\right] \mu_m\right|\right|_\infty = C \left|\left|(K^m - K_A^m)\, \mu_m\right|\right|_\infty \tag{3.42}$$

holds for sufficiently large A and some constant C. Since

$$\left|\left|(K^m - K_A^m)\, \mu_m\right|\right|_\infty \leq \left|\left|(K^m - K_A^m)\right|\right|_\infty \left|\left|\mu_m\right|\right|_\infty \to 0 \tag{3.43}$$

as $A \to \infty$, we conclude that the solutions $\mu_m^A(x)$ converge in norm (i.e., converge uniformly for $x \in [0, L]$) to $\mu_m(x)$ as $A \to \infty$ for k which are not Wood Anomaly values. □

3.1.2 Regularity of Solutions for Smooth Gratings

In this section, we apply the regularity theory presented in [43]—much of which is derived from the results of [46]. For problems having grating profiles $f(x)$ belonging to $C^\infty_{\text{per}}(L)$ and for k which are not Wood Anomaly values, we show that the solutions $\mu_m \in C^0_{\text{per}}(L)$ of the exact integral equations (3.5) can be represented as Fourier series of the form

$$\sum_{n=-\infty}^{\infty} d_n e^{i\frac{2\pi n}{L}x}$$

which converge absolutely and uniformly, where the magnitudes of the coefficients d_n decrease super-algebraically in n as $n \to \pm\infty$; equivalently, $\mu_m \in C^\infty_{\text{per}}(L)$. This regularity property will lead us to conclude in the following section that the solutions $\mu_m^A(x)$ of the approximating integral equations (3.1) converge super-algebraically in A to $\mu_m(x)$ as $A \to \infty$ for problems with smooth gratings. Additionally, we show that in such cases $\mu_m^A(x)$ belong to $C^\infty_{\text{per}}(L)$ as well, which will allow us to demonstrate in Section 3.2.1.3 that the corresponding numerical solutions that are computed by our method are super-algebraically convergent in the number of discretization points.

Remark 3.1.2. *The regularity of $\mu_m(x)$ and $\mu_m^A(x)$ for finitely differentiable profiles $f(x)$ will not be discussed here; see [43] for a theoretical discussion in these regards. Such analysis can be developed by using minor modifications of the reasoning presented in this section.*

We begin by establishing certain notation for this section that is in keeping with [43].

Let

$$\Omega^+ \equiv \{\boldsymbol{r} : -\gamma < x < L+\gamma,\ f(x) < y < H_+\}, \quad \Omega^- \equiv \{\boldsymbol{r} : -\gamma < x < L+\gamma,\ H_- < y < f(x)\} \tag{3.44}$$

for some real $\gamma > 0$, $H_+ > \max[f(x)]$ and $H_- < \min[f(x)]$, and let

$$\Gamma \equiv \partial\Omega^+. \tag{3.45}$$

Also, let G_1 and G_2 be bounded open subsets of $\mathbb{R}^2$ such that $\bar{G_1}$ is a compact subset of G_2; G_1 and G_2 each intersect Γ, but each intersection is only on an open subset of $\boldsymbol{r}$ for which $y = f(x)$. Define

$$\Omega_j^\pm \equiv G_j \cap \Omega^\pm, \quad \Gamma_j \equiv G_j \cap \Gamma \tag{3.46}$$

for $j = 1, 2$. For integer $\ell \geq 0$ and real $0 < \alpha \leq 1$, we say that the scattering surface $y = f(x)$ is $C^{\ell,\alpha}(\mathbb{R})$ if there exists a constant C such that

$$\left| f^{(\ell)}(x) - f^{(\ell)}(x') \right| \leq C|x - x'|^\alpha \tag{3.47}$$

for every $x, x' \in \mathbb{R}$, i.e., if the ℓth derivative of f is uniformly Hölder continuous with Hölder exponent α. Finally, for real s, a function $u(\boldsymbol{r})$ belongs to $H^s\left(\mathbb{R}^2\right)$—the Sobolev space of order s on $\mathbb{R}^2$—if $\mathcal{J}^s u \in L_2\left(\mathbb{R}^2\right)$, where

$$\mathcal{J}^s u(\boldsymbol{r}) \equiv \int_{\mathbb{R}^2} \left(1 + |\boldsymbol{\xi}|^2\right)^{\frac{s}{2}} \hat{u}(\boldsymbol{\xi}) e^{i2\pi\boldsymbol{\xi}\cdot\boldsymbol{r}}\, d\boldsymbol{\xi}, \tag{3.48}$$

$$\hat{u}(\boldsymbol{\xi}) \equiv \int_{\mathbb{R}^2} u(\boldsymbol{r}) e^{-i2\pi\boldsymbol{\xi}\cdot\boldsymbol{r}}\, d\boldsymbol{r} \tag{3.49}$$

(i.e., $\hat{u}(\boldsymbol{\xi})$ is the Fourier transform of $u(\boldsymbol{r})$) and

$$L_2\left(\mathbb{R}^2\right) \equiv \left\{ \phi : \int_{\mathbb{R}^2} |\phi(\boldsymbol{r})|^2\, d\boldsymbol{r} < \infty \right\} \tag{3.50}$$

is the space of square-integrable functions on $\mathbb{R}^2$. The spaces $H^s\left(\Omega_j^\pm\right)$ and $H^s\left(\Gamma_j\right)$ are defined with respect to $H^s\left(\mathbb{R}^2\right)$; see [43, pp. 77, 98, 99] for details.

We now prove certain intermediate theorems for the TE/sound-soft and TM/sound-hard

scattering problems.

Theorem 3.1.5. *Let the grating profile $f(x)$ belong to $C^{\infty}_{per}(L)$. Then, for k which are not Wood Anomaly values, the solution $\mu(\boldsymbol{r})$ of the TE/sound-soft integral equation (2.35) belongs to $H^{r+\frac{1}{2}}(\Gamma_1)$ for every integer $r \geq 0$.*

Proof. As is well known (see, e.g., [19]), for k which are not Wood Anomaly values, the TE/sound-soft scattered field (formulated in Section 2.1.3.1)

$$\psi^{\text{scat}}(\boldsymbol{r}) = -\int_{\mathcal{P}(x)} \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')\mu(\boldsymbol{r}')\, ds(\boldsymbol{r}') \tag{3.51}$$

is continuous throughout $\mathbb{R}^2$, and its normal derivatives on Γ_2—taken as limits from above and below the profile—satisfy the jump condition

$$\frac{\partial \psi^{\text{scat}}_{+}(\boldsymbol{r})}{\partial \boldsymbol{\nu}(\boldsymbol{r})} - \frac{\partial \psi^{\text{scat}}_{-}(\boldsymbol{r})}{\partial \boldsymbol{\nu}(\boldsymbol{r})} = \mu(\boldsymbol{r}), \quad \boldsymbol{r} \in \Gamma_2, \tag{3.52}$$

where

$$\frac{\partial \psi^{\text{scat}}_{\pm}(\boldsymbol{r})}{\partial \boldsymbol{\nu}(\boldsymbol{r})} \equiv \lim_{\epsilon \to +0} \nabla \psi^{\text{scat}}(\boldsymbol{r} \pm \epsilon \boldsymbol{\nu}(\boldsymbol{r})) \cdot \boldsymbol{\nu}(\boldsymbol{r}), \quad \boldsymbol{r} \in \Gamma_2. \tag{3.53}$$

The field $\psi^{\text{scat}}(\boldsymbol{r})$ satisfies the Helmholtz equation (2.6) on $\Omega_2^{\pm}$, and the density $\mu(\boldsymbol{r})$ is not only a solution of (2.35), but (as discussed in Remark 2.1.6) it also satisfies the first kind integral equation

$$\int_{\mathcal{P}(x)} \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')\mu(\boldsymbol{r}')\, ds(\boldsymbol{r}') = \psi^{\text{inc}}(\boldsymbol{r}) = e^{ik(\sin(\theta), -\cos(\theta))\cdot \boldsymbol{r}}, \quad \boldsymbol{r} \in \Gamma_2 \tag{3.54}$$

because of the Dirichlet boundary condition

$$\psi^{\text{scat}}(\boldsymbol{r}) = -\psi^{\text{inc}}(\boldsymbol{r}), \quad \boldsymbol{r} \in \Gamma_2. \tag{3.55}$$

Thus, since $f \in C^{\infty}_{\text{per}}(L)$, we have $\Gamma_2 \in C^{r+1,1}$ and $\psi^{\text{inc}} \in H^{r+\frac{3}{2}}(\Gamma_2)$ for every integer $r \geq 0$, and there exists some $\mu \in H^{-\frac{1}{2}}(\Gamma)$ (in fact, $\mu(\boldsymbol{r})$ is continuous) such that (3.54) is satisfied on Γ_2. Therefore, we invoke [43, Theorem 7.16(i)] and obtain the following result: $\mu \in H^{r+\frac{1}{2}}(\Gamma_1)$ for every integer $r \geq 0$. □

Theorem 3.1.6. *Let the grating profile $f(x)$ belong to $C^{\infty}_{per}(L)$. Then, for k which are not Wood Anomaly values, the solution $\mu(\boldsymbol{r})$ of the TM/sound-hard integral equation (2.39)*

belongs to $H^{r+\frac{1}{2}}(\Gamma_1)$ *for every integer* $r \geq 0$.

Proof. For k which are not Wood Anomaly values, the TM/sound-hard scattered field

$$\psi^{\text{scat}}(\boldsymbol{r}) = \int_{\mathcal{P}(x)} \Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')\mu(\boldsymbol{r}')\, ds(\boldsymbol{r}') \tag{3.56}$$

is continuous throughout $\mathbb{R}^2$ and satisfies both the Helmholtz equation on $\Omega_2^{\pm}$ and the jump condition

$$\frac{\partial \psi_+^{\text{scat}}(\boldsymbol{r})}{\partial \boldsymbol{\nu}(\boldsymbol{r})} - \frac{\partial \psi_-^{\text{scat}}(\boldsymbol{r})}{\partial \boldsymbol{\nu}(\boldsymbol{r})} = -\mu(\boldsymbol{r}), \quad \boldsymbol{r} \in \Gamma_2. \tag{3.57}$$

By (3.56), we have that $\psi^{\text{scat}} \in C^{\infty}(\Omega_2^+)$ since $\Phi_{\text{per}}(\boldsymbol{r}, \boldsymbol{r}')$ (2.21) is $C^{\infty}(\Omega_2^+)$. Thus, for every integer $r \geq 0$, $\psi^{\text{scat}} \in H^{r+2}(\Omega_2^+)$ and $\Gamma_2 \in C^{r+1,1}$; it follows that $\psi^{\text{scat}} \in H^{r+\frac{3}{2}}(\Gamma_2)$ [43, Theorem 3.37]. Because (3.56) is satisfied on Γ_2, the desired result again follows by [43, Theorem 7.16(i)]. □

For a function $\phi \in L^2[0, L]$, let

$$\sum_{n=-\infty}^{\infty} d_n e^{i\frac{2\pi n}{L}x}$$

be the Fourier series of ϕ. Then, consistent with our earlier notation and also in keeping with [35, Chapter 8], we let $H^s[0, L]$ be the Sobolev space of order s on the interval $[0, L]$. For any real $s \geq 0$, this is the space of all functions $\phi \in L^2[0, L]$ whose Fourier coefficients d_n satisfy

$$\sum_{n=-\infty}^{\infty} \left(1 + n^2\right)^s |d_n|^2 < \infty. \tag{3.58}$$

We note that for $s > \frac{1}{2}$ the Fourier series for ϕ converges absolutely and uniformly, and its limit belongs to $C^0_{\text{per}}(L)$ and coincides with ϕ almost everywhere [35, Theorem 8.4]; if $\phi \in C^0_{\text{per}}(L) \cap H^s[0, L]$ for some real $s > \frac{1}{2}$, then the Fourier series for $\phi(x)$ converges absolutely and uniformly to $\phi(x)$ for $x \in [0, L]$.

Therefore, the following result holds.

Theorem 3.1.7. *Let the grating profile* $f(x)$ *belong to* $C^{\infty}_{per}(L)$*, and let* k *be such that it is not a Wood Anomaly value. Then, the solutions* $\mu_m \in C^0_{per}(L)$ $(m = 1, 2)$ *of the exact integral equations* (3.5) *belong to* $H^{r+\frac{1}{2}}[0, L]$ *for every integer* $r \geq 0$*; in particular, they can*

be represented as Fourier series of the form

$$\sum_{n=-\infty}^{\infty} d_n e^{i\frac{2\pi n}{L}x}$$

which converge absolutely and uniformly, where the magnitudes of the coefficients d_n decrease super-algebraically in n as $n \to \pm\infty$. Equivalently, $\mu_m \in C^{\infty}_{per}(L)$.

Proof. As stated in Sections 2.3.1 and 2.3.2.2, the solutions $\mu_m \in C^0_{\text{per}}(L)$ of (3.5) are related to the solution $\mu(\boldsymbol{r})$ of either (2.35) (TE/sound-soft scattering) or (2.39) (TM/sound-hard scattering) by the formulas

$$\mu_1(x) \equiv \mu(\boldsymbol{r})\frac{e^{-ik\sin(\theta)x}}{-2ki}\sqrt{1+[f'(x)]^2} \tag{3.59}$$

and

$$\mu_2(x) \equiv \mu(\boldsymbol{r})\frac{e^{-ik\sin(\theta)x+ik\cos(\theta)f(x)}}{-2ki}\sqrt{1+[f'(x)]^2}. \tag{3.60}$$

We choose the subsets G_1 and G_2 so that Γ_1 contains $\{\boldsymbol{r} : 0 \le x \le L,\ y = f(x)\}$. Since $f \in C^{\infty}_{\text{per}}(L)$, it follows from Theorems 3.1.5 and 3.1.6 and the above formulas that $\mu_m \in H^{r+\frac{1}{2}}[0,L]$ for every integer $r \ge 0$. Thus, the Fourier series expansions of $\mu_m(x)$ converge absolutely and uniformly to $\mu_m(x)$, and it follows by the property (3.58) that the magnitudes of the Fourier coefficients d_n decrease super-algebraically in n as $n \to \pm\infty$. Also, $\mu_m \in C^0_{\text{per}}(L) \cap H^{r+\frac{1}{2}}[0,L]$ for every integer $r \ge 0$ implies that $\mu_m \in C^{\infty}_{\text{per}}(L)$. □

Finally, we conclude this section with the corresponding theorem for $\mu^A_m(x)$.

Theorem 3.1.8. *Let the grating profile $f(x)$ belong to $C^{\infty}_{per}(L)$, and let k be such that it is not a Wood Anomaly value. Then, the solutions $\mu^A_m \in C^0_{per}(L)$ $(m = 1, 2)$ of the approximating integral equations (3.1) belong to $H^{r+\frac{1}{2}}[0,L]$ for every integer $r \ge 0$; equivalently, $\mu^A_m \in C^{\infty}_{per}(L)$.*

Proof. Using the formulas

$$\mu^A(\boldsymbol{r}) \equiv \mu^A_1(x)\frac{-2kie^{ik\sin(\theta)x}}{\sqrt{1+[f'(x)]^2}} = \mu^A_2(x)\frac{-2kie^{ik\sin(\theta)x-ik\cos(\theta)f(x)}}{\sqrt{1+[f'(x)]^2}}, \tag{3.61}$$

we re-write (3.1) as

$$\frac{1}{2}\mu^A(\boldsymbol{r})\pm\int_{x-A}^{x+A} P_1(x,x',c,A)\frac{i}{4}\frac{\partial H_0^1(ku(x,x'))}{\partial\boldsymbol{\nu}(\boldsymbol{r})}\mu^A(\boldsymbol{r}')\sqrt{1+[f'(x')]^2}\,dx' = \frac{\partial\psi^{\text{inc}}(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})}, \quad y=f(x). \tag{3.62}$$

Given (2.55), this can be re-expressed as

$$\begin{aligned}\frac{1}{2}\mu^A(\boldsymbol{r}) &\pm \int_{-\infty}^{\infty}\frac{i}{4}\frac{\partial H_0^1(ku(x,x'))}{\partial\boldsymbol{\nu}(\boldsymbol{r})}\mu^A(\boldsymbol{r}')\sqrt{1+[f'(x')]^2}\,dx' \\ &= \frac{\partial\psi^{\text{inc}}(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})} \pm \int_{-\infty}^{\infty} P_2(x,x',c,A)\frac{i}{4}\frac{\partial H_0^1(ku(x,x'))}{\partial\boldsymbol{\nu}(\boldsymbol{r})}\mu^A(\boldsymbol{r}')\sqrt{1+[f'(x')]^2}\,dx', \quad y=f(x).\end{aligned} \tag{3.63}$$

For both the TE/sound-soft and TM/sound-hard problems, we define

$$\psi^A(\boldsymbol{r}) = \int_{\mathcal{P}(x)} \Phi_{\text{per}}(\boldsymbol{r},\boldsymbol{r}')\mu^A(\boldsymbol{r}')\,ds(\boldsymbol{r}'). \tag{3.64}$$

Like the function $\psi^{\text{scat}}(\boldsymbol{r})$ (3.56) in Theorem 3.1.6, $\psi^A(\boldsymbol{r})$ is continuous throughout $\mathbb{R}^2$ and satisfies both the Helmholtz equation on $\Omega_2^{\pm}$ and the jump condition

$$\frac{\partial\psi_+^A(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})} - \frac{\partial\psi_-^A(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})} = -\mu^A(\boldsymbol{r}), \quad \boldsymbol{r}\in\Gamma_2. \tag{3.65}$$

Also, we note that the right-hand side of (3.63) is C^∞, since both ψ^{inc} and $P_2(x,x',c,A)\frac{\partial H_0^1(ku(x,x'))}{\partial\boldsymbol{\nu}(\boldsymbol{r})}$ are smooth functions of x due to the smoothness of the profile $f(x)$ and the fact that $P_2(x,x',c,A)=0$ in a neighborhood of $x'=x$.

Considering first the TE/sound-soft problem, we have

$$\begin{aligned}\frac{\partial\psi_-^A(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})} &= \frac{1}{2}\mu^A(\boldsymbol{r}) + \int_{\mathcal{P}(x)}\frac{\partial\Phi_{\text{per}}(\boldsymbol{r},\boldsymbol{r}')}{\partial\boldsymbol{\nu}(\boldsymbol{r})}\mu^A(\boldsymbol{r}')\,ds(\boldsymbol{r}') \\ &= \frac{1}{2}\mu^A(\boldsymbol{r}) + \int_{-\infty}^{\infty}\frac{i}{4}\frac{\partial H_0^1(ku(x,x'))}{\partial\boldsymbol{\nu}(\boldsymbol{r})}\mu^A(\boldsymbol{r}')\sqrt{1+[f'(x')]^2}\,dx'.\end{aligned} \tag{3.66}$$

Since the right-hand side of (3.63) is C^∞, it follows that $\psi^A(\boldsymbol{r})\in H^{r+\frac{3}{2}}(\Gamma_2)$ for every integer $r\geq 0$. Therefore, $\mu^A(\boldsymbol{r})\in H^{r+\frac{1}{2}}(\Gamma_1)$ for every integer $r\geq 0$ [43, Theorem 7.16(i)], and we conclude that $\mu_m^A\in C_{\text{per}}^\infty(L)$.

The TM/sound-hard problem is handled similarly:

$$\begin{aligned}\frac{\partial\psi_+^A(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})} &= -\frac{1}{2}\mu^A(\boldsymbol{r}) + \int_{\mathcal{P}(x)} \frac{\partial\Phi_{\text{per}}(\boldsymbol{r},\boldsymbol{r}')}{\partial\boldsymbol{\nu}(\boldsymbol{r})}\mu^A(\boldsymbol{r}')\,ds(\boldsymbol{r}') \\ &= -\frac{1}{2}\mu^A(\boldsymbol{r}) + \int_{-\infty}^{\infty} \frac{i}{4}\frac{\partial H_0^1(ku(x,x'))}{\partial\boldsymbol{\nu}(\boldsymbol{r})}\mu^A(\boldsymbol{r}')\sqrt{1+[f'(x')]^2}\,dx',\end{aligned} \tag{3.67}$$

so we have

$$\frac{\partial\psi_+^A(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})} = -\frac{\partial\psi^{\text{inc}}(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})} + \int_{-\infty}^{\infty} P_2(x,x',c,A)\frac{i}{4}\frac{\partial H_0^1(ku(x,x'))}{\partial\boldsymbol{\nu}(\boldsymbol{r})}\mu^A(\boldsymbol{r}')\sqrt{1+[f'(x')]^2}\,dx', \quad y=f(x). \tag{3.68}$$

Again, it follows that $\mu_m^A \in C_{\text{per}}^\infty(L)$. □

3.1.3 Super-Algebraic Convergence of the Integral Operators and Solutions for Smooth Gratings

Earlier (Section 2.2.3), we demonstrated that a highly simplified version of the integral operator in the scattering equations (2.51) and (2.52) can be approximated well by using the C^∞ windowing function $P_1(x,x',c,A)$. In particular, the error of approximation for this example decreases super-algebraically in A as $A\to\infty$, even though the integrand (without the factor $P_1(x,x',c,A)$) is $\mathcal{O}\left(\frac{1}{\sqrt{x'}}\right)$ as $x'\to\infty$; this is due to the oscillatory component of the integrand along with the smooth decay of $P_1(x,x',c,A)$ to 0, as was shown via integration by parts. In this section, we employ certain series expansions and once again use integration by parts in order to establish with the full complexities of the problem—without any simplifications—that the approximations of the integral operators of our method similarly converge super-algebraically as $A\to\infty$ for scattering configurations with smooth grating profiles. In view of the error estimate (3.42), this implies that the solutions $\mu_m^A(x)$ converge super-algebraically in A, uniformly for $x\in[0,L]$, to $\mu_m(x)$ as $A\to\infty$ for k which are not Wood Anomaly values.

We recall that the operator K_A^m, given by

$$\begin{aligned}K_A^m\varphi(x) &= \int_{x-A}^{x+A} P_1(x,x',c,A)g(x,x')\frac{h(ku(x,x'))}{e^{iku(x,x')}}e^{ik\phi_m(x,x')}\varphi(x')\,dx' \\ &= \int_0^L G_A^m(x,x',k,c)\varphi(x')\,dx', \quad 0\le x\le L,\end{aligned} \tag{3.69}$$

converges in the $C^0_{\text{per}}(L)$ norm to K^m, given by

$$\begin{aligned} K^m\varphi(x) &= \int_{-\infty}^{\infty} g(x,x')\frac{h(ku(x,x'))}{e^{iku(x,x')}}e^{ik\phi_m(x,x')}\varphi(x')\,dx' \\ &= \int_0^L G^m(x,x',k)\varphi(x')\,dx', \quad 0 \le x \le L, \end{aligned} \tag{3.70}$$

as $A \to \infty$ for k which are not Wood Anomaly values, i.e., that

$$\begin{aligned} ||K^m - K^m_A||_\infty &= \sup_{\varphi \in C^0_{\text{per}}(L)} \frac{||(K^m - K^m_A)\,\varphi||_\infty}{||\varphi||_\infty} \\ &= \sup_{\varphi \in C^0_{\text{per}}(L)} \frac{\max_{x\in[0,L]} |(K^m - K^m_A)\,\varphi(x)|}{\max_{x\in[0,L]} |\varphi(x)|} \\ &\to 0 \end{aligned} \tag{3.71}$$

as $A \to \infty$ for such k (Section 3.1.1.1). Keeping in mind that the solutions $\mu_m(x)$ of the full integral equations (3.5) belong to $C^\infty_{\text{per}}(L)$ if the scattering surface $f(x)$ belongs to $C^\infty_{\text{per}}(L)$ (Section 3.1.2), we here show that

$$|(K^m - K^m_A)\,\varphi(x)| = \left| \int_{-\infty}^{\infty} P_2(x,x',c,A)g(x,x')\frac{h(ku(x,x'))}{e^{iku(x,x')}}e^{ik\phi_m(x,x')}\varphi(x')\,dx' \right| \tag{3.72}$$

decreases super-algebraically in A as $A \to \infty$, uniformly for all $x \in [0,L]$, for $f, \varphi \in C^\infty_{\text{per}}(L)$ and k which are not Wood Anomaly values. More precisely, let

$$\begin{aligned} (K^m - K^m_A)\,\varphi(x) &\equiv I^m(x,k,c,A) \\ &= I^m_+(x,k,c,A) + I^m_-(x,k,c,A), \end{aligned} \tag{3.73}$$

where

$$\begin{aligned} &I^m_+(x,k,c,A) \\ &\quad \equiv \int_x^{\infty} P_2(x,x',c,A)g(x,x')\frac{h(ku(x,x'))}{e^{iku(x,x')}}e^{ik\phi_m(x,x')}\varphi(x')\,dx' \\ &\quad = \int_0^{\infty} P_2(x,x+x',c,A)g(x,x+x')\frac{h(ku(x,x+x'))}{e^{iku(x,x+x')}}e^{ik\phi_m(x,x+x')}\varphi(x+x')\,dx' \end{aligned} \tag{3.74}$$

and

$$\begin{aligned} &I_-^m(x,k,c,A) \\ &\quad\equiv \int_{-\infty}^{x} P_2(x,x',c,A)g(x,x')\frac{h(ku(x,x'))}{e^{iku(x,x')}}e^{ik\phi_m(x,x')}\varphi(x')\,dx' \\ &\quad= \int_{-\infty}^{0} P_2(x,x+x',c,A)g(x,x+x')\frac{h(ku(x,x+x'))}{e^{iku(x,x+x')}}e^{ik\phi_m(x,x+x')}\varphi(x+x')\,dx'. \end{aligned} \tag{3.75}$$

Then, the following theorem holds.

Theorem 3.1.9. *Let the grating profile $f(x)$ belong to $C^\infty_{per}(L)$, and let $\varphi(x)$ belong to $C^\infty_{per}(L)$ as well. Also, let η be a real number such that $\eta > 0$. Given $I_+^m(x,k,c,A)$ (3.74) and $I_-^m(x,k,c,A)$ (3.75),*

1. *if for every $n \in \mathbb{Z}$*

$$|kL\,[1+\sin(\theta)] + 2\pi n| > \eta > 0, \tag{3.76}$$

then

$$\max_{x\in[0,L]} \left|I_+^m(x,k,c,A)\right| = \mathcal{O}\left(\left(\frac{A}{L}\right)^{-p+\frac{1}{2}}\right) \tag{3.77}$$

as $A \to \infty$ for every integer $p \geq 1$;

2. *if for every $n \in \mathbb{Z}$*

$$|kL\,[1-\sin(\theta)] + 2\pi n| > \eta > 0, \tag{3.78}$$

then

$$\max_{x\in[0,L]} \left|I_-^m(x,k,c,A)\right| = \mathcal{O}\left(\left(\frac{A}{L}\right)^{-p+\frac{1}{2}}\right) \tag{3.79}$$

as $A \to \infty$ for every integer $p \geq 1$;

3. *if for every $n \in \mathbb{Z}$*

$$|kL\,[1\pm\sin(\theta)] + 2\pi n| > \eta > 0, \tag{3.80}$$

that is, if k is some distance away from all Wood Anomaly values, then

$$\begin{aligned} \max_{x\in[0,L]} \left|\left(K^m - K_A^m\right)\varphi(x)\right| &= \max_{x\in[0,L]} \left|I^m(x,k,c,A)\right| \\ &= \mathcal{O}\left(\left(\frac{A}{L}\right)^{-p+\frac{1}{2}}\right) \end{aligned} \tag{3.81}$$

as $A \to \infty$ for every integer $p \geq 1$.

Remark 3.1.3. *Some of the intermediate results found in the following sections allow for estimates of these integrals to be obtained given finitely differentiable functions $f(x)$ and $\varphi(x)$. In particular, the sizes of the integrals would be algebraically small in $\frac{A}{L}$ as $A \to \infty$ rather than super-algebraically small.*

After a discussion of certain changes of variables and series expansions in Section 3.1.3.1 and a verification of several lemmas in Section 3.1.3.2, we prove Theorem 3.1.9 in Section 3.1.3.3. Complete details are provided only for $I_+^2(x,k,c,A)$, since the proofs relating to $I_-^2(x,k,c,A)$, $I_+^1(x,k,c,A)$ and $I_-^1(x,k,c,A)$ are very similar. This leads to our super-algebraic convergence result for the solutions $\mu_m^A(x)$ (Section 3.1.3.4).

3.1.3.1 Preliminary Considerations

We begin by establishing a number of preliminary results that will facilitate our analysis of $I_+^2(x,k,c,A)$. A similar discussion for $I_-^2(x,k,c,A)$ will be provided later.

Since $P_2(x,x+x',c,A) = 0$ for $x' \in [0,cA]$, we have

$$I_+^2(x,k,c,A) = \int_{cA}^{\infty} P_2(x,x+x',c,A) g(x,x+x') \frac{h(ku(x,x+x'))}{e^{iku(x,x+x')}} e^{ik\phi_2(x,x+x')} \varphi(x+x')\, dx'. \tag{3.82}$$

To treat this integral, we non-dimensionalize its variables and functions by employing L (the period of $f(x)$ and $\varphi(x)$) and h (the height of $f(x)$). Using the change of variables $t \equiv \frac{x'}{L}$ and defining

$$\frac{h}{2}\bar{f}\left(\frac{2\pi x}{L}\right) \equiv f(x) \longrightarrow f'(x) = \frac{h}{2L} 2\pi \bar{f}'\left(\frac{2\pi x}{L}\right) \tag{3.83}$$

and

$$\bar{\varphi}\left[2\pi\left(\frac{x}{L}+t\right)\right] \equiv \varphi(x+Lt), \tag{3.84}$$

we obtain

$$I_+^2(x,k,c,A) = \int_{\frac{cA}{L}}^{\infty} \bar{P}_2\left(c, \frac{Lt}{cA}\right) \bar{g}_+(x,t) \frac{h\left[kL\bar{u}_+(x,t)\right]}{e^{ikL\bar{u}_+(x,t)}} e^{ikL\bar{\phi}_+(x,t)} \bar{\varphi}\left[2\pi\left(\frac{x}{L}+t\right)\right] dt. \tag{3.85}$$

Here, taking into account (2.47), (2.53)–(2.55), (2.85) and (2.103), we have set

$$\bar{P}_2\left(c, \frac{Lt}{cA}\right) \equiv 1 - S\left(\frac{Lt}{cA}, 1, \frac{1}{c}\right), \tag{3.86}$$

$$\bar{g}_+(x,t) \equiv \frac{i}{2}\frac{h}{2L}\frac{1}{t}\frac{\frac{\bar{f}\left[2\pi\left(\frac{x}{L}+t\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)}{t} - 2\pi\bar{f}'\left(\frac{2\pi x}{L}\right)}{1+\left(\frac{h}{2L}\right)^2\left\{\frac{\bar{f}\left[2\pi\left(\frac{x}{L}+t\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)}{t}\right\}^2}, \tag{3.87}$$

$$\bar{u}_+(x,t) \equiv t\sqrt{1+\left(\frac{h}{2L}\right)^2\left\{\frac{\bar{f}\left[2\pi\left(\frac{x}{L}+t\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)}{t}\right\}^2} \tag{3.88}$$

and

$$\begin{aligned}\bar{\phi}_+(x,t) \equiv\ & t\sqrt{1+\left(\frac{h}{2L}\right)^2\left\{\frac{\bar{f}\left[2\pi\left(\frac{x}{L}+t\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)}{t}\right\}^2} \\ & + (\sin(\theta),-\cos(\theta))\cdot\left(t,\frac{h}{2L}\left\{\bar{f}\left[2\pi\left(\frac{x}{L}+t\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)\right\}\right).\end{aligned} \tag{3.89}$$

In view of (2.105), $\frac{\partial\bar{\phi}_+}{\partial t}$ has an upper and lower bound for $t\in(0,\infty)$. Let $c_1\geq 0$ be a constant such that

$$\frac{\partial}{\partial t}\left[\bar{\phi}_+(x,t)+c_1 t\right] > \epsilon > 0 \tag{3.90}$$

for all $(x,t)\in[0,L]\times(0,\infty)$. This equation implies that, for each $x\in[0,L]$, $\bar{\phi}_+(x,t)+c_1 t$ has a bounded inverse for all $t\in\left[\frac{cA}{L},\infty\right)$. Using a second change of variables

$$\xi \equiv \tilde{\phi}_+(x,t), \tag{3.91}$$

where

$$\tilde{\phi}_+(x,t) \equiv \frac{1}{1+\sin(\theta)+c_1}\left[\bar{\phi}_+(x,t)+c_1 t\right] \tag{3.92}$$

($2>1+\sin(\theta)>\epsilon_0>0$ for $-\frac{\pi}{2}+\delta_0<\theta<\frac{\pi}{2}-\delta_0$, which implies that $\frac{1}{\epsilon_0+c_1}>\frac{1}{1+\sin(\theta)+c_1}>\frac{1}{2+c_1}>0$) and

$$\tilde{\phi}_+^{-1}(x,\xi) \equiv t, \tag{3.93}$$

we thus have

$$\begin{aligned}I_+^2(x,k,c,A) = \int_{\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty}& \bar{P}_2\left(c,\frac{L\tilde{\phi}_+^{-1}}{cA}\right)\bar{g}_+\left(x,\tilde{\phi}_+^{-1}\right)\frac{h\left(kL\bar{u}_+\left(x,\tilde{\phi}_+^{-1}\right)\right)}{e^{ikL\bar{u}_+\left(x,\tilde{\phi}_+^{-1}\right)}} \\ & \times e^{ikL[1+\sin(\theta)+c_1]\xi}e^{-ikLc_1\tilde{\phi}_+^{-1}}\bar{\varphi}\left[2\pi\left(\frac{x}{L}+\tilde{\phi}_+^{-1}\right)\right]\frac{\partial\tilde{\phi}_+^{-1}}{\partial\xi}\,d\xi.\end{aligned} \tag{3.94}$$

Remark 3.1.4. *We may let $c_1=0$ for simple-reflection cases, since for such problems we*

have $\frac{\partial\phi(x,x')}{\partial x'} > 0$ *for all* $x' > x$ *(Corollary 2.3.1). But, if multiple reflections are present, then we must choose* $c_1 > 0$ *so that* (3.90) *is satisfied.*

Now, since $\varphi \in C^{\infty}_{\text{per}}(L)$, we may write

$$\bar{\varphi}\left[2\pi\left(\frac{x}{L}+t\right)\right] = \varphi(x+Lt) = \sum_{n=-\infty}^{\infty} d_n e^{i\frac{2\pi n}{L}(x+Lt)}; \tag{3.95}$$

this Fourier series converges absolutely and uniformly to $\bar{\varphi}\left[2\pi\left(\frac{x}{L}+t\right)\right]$ for $(x,t) \in [0,L] \times \left[\frac{cA}{L},\infty\right)$, and the magnitudes of the coefficients d_n decrease super-algebraically in n as $n \to \pm\infty$ (Theorem 3.1.7), i.e.,

$$|d_n| \leq C_s |n|^{-s} \tag{3.96}$$

for all $n \in \mathbb{Z}$ for every integer $s \geq 1$ and certain constants C_s that are independent of n. Substituting (3.93) together with (3.95) into (3.94), by uniform convergence of the series we obtain

$$I^2_+(x,k,c,A) = \sum_{n=-\infty}^{\infty} I_n(x,k,c,A), \tag{3.97}$$

where

$$\begin{aligned} I_n(x,k,c,A) \equiv d_n e^{in\frac{2\pi x}{L}} \int_{\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty} \bar{P}_2\left(c, \frac{L\tilde{\phi}_+^{-1}}{cA}\right) \bar{g}_+\left(x,\tilde{\phi}_+^{-1}\right) \frac{h\left[kL\bar{u}_+\left(x,\tilde{\phi}_+^{-1}\right)\right]}{e^{ikL\bar{u}_+\left(x,\tilde{\phi}_+^{-1}\right)}} \\ \times e^{ikL[1+\sin(\theta)+c_1]\xi} e^{i(2\pi n - kLc_1)\tilde{\phi}_+^{-1}} \frac{\partial\tilde{\phi}_+^{-1}}{\partial\xi}\, d\xi. \end{aligned} \tag{3.98}$$

As we will show, the series (3.97) is absolutely convergent, so we therefore estimate the size of $|I_n(x,k,c,A)|$ as $A \to \infty$ in order to prove Theorem 3.1.9.

We establish a number of lemmas in Section 3.1.3.2 through the use of Taylor's formula [3]

$$s(x+a) = s(x) + as'(x) + \frac{a^2}{2!}s''(x) + \ldots + \frac{a^{\gamma-1}}{(\gamma-1)!}s^{(\gamma-1)}(x) + \frac{a^{\gamma}}{(\gamma-1)!}\int_0^1 (1-y)^{\gamma-1} s^{(\gamma)}(x+ya)\, dy \tag{3.99}$$

for $s \in C^{\gamma-1}[x,x+a] \cup C^{\gamma}(x,x+a)$, and this will allow us in Section 3.1.3.3 to produce a useful expansion of the integrand of each $I_n(x,k,c,A)$. More precisely, we will write this integrand as a sum of terms with periodic and/or smoothly decaying factors, and, having this expansion, we will prove Theorem 3.1.9 via repeated integrations by parts which involve

integrating the periodic factors and differentiating the smoothly decaying factors.

3.1.3.2 Foundational Lemmas

Recalling the formula (2.1) for $C^r_{\text{per}}(L)$, it is convenient for us here to introduce the additional notation

$$C^r_{\text{bdd}}[b,\infty) \equiv \{f \in C^r[b,\infty) : f \text{ is bounded for } [b,\infty)\}. \tag{3.100}$$

We now establish a sequence of important lemmas that we will use in the next section.

Lemma 3.1.3. *For $t > 0$, let the functions $f(t)$ and $g(t)$ be given by*

$$f(t) \equiv \sum_{n=0}^{p-1} \frac{S_n(t)}{t^n} + \frac{S_p^{rem}(t)}{t^p} \tag{3.101}$$

and

$$g(t) \equiv \sum_{m=0}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{rem}(t)}{t^q}, \tag{3.102}$$

where $q \geq p \geq 1$. The function $f(t)$ has the properties $S_n \in C^{r_f-n}_{per}(L)$ and $S_p^{rem} \in C^{r_f-p}_{bdd}[b,\infty)$ for some integer $r_f \geq p$ and some real $b > 0$; $g(t)$ has the properties $T_m \in C^{r_g-m}_{per}(L)$ and $T_q^{rem} \in C^{r_g-q}_{bdd}[b,\infty)$ for some integer $r_g \geq q$. Then, $f(t)g(t)$ admits a similar representation:

$$f(t)g(t) = \sum_{\ell=0}^{p-1} \frac{W_\ell(t)}{t^\ell} + \frac{W_p^{rem}(t)}{t^p}, \tag{3.103}$$

where $W_\ell \in C^{\min(r_f,r_g)-\ell}_{per}(L)$ and $W_p^{rem} \in C^{\min(r_f-p,r_g-q)}_{bdd}[b,\infty)$.

Proof. The properties of $f(t)$ and $g(t)$ imply that $S_n T_{\ell-n} \in C^{\min(r_f-n,r_g-\ell+n)}_{\text{per}}(L)$ and

$S_p^{\rm rem}T_q^{\rm rem} \in C_{\rm bdd}^{\min(r_f-p,r_g-q)}[b,\infty)$. Thus,

$$\begin{aligned}
f(t)g(t) &= \left[\sum_{n=0}^{p-1}\frac{S_n(t)}{t^n}+\frac{S_p^{\rm rem}(t)}{t^p}\right]\left[\sum_{m=0}^{q-1}\frac{T_m(t)}{t^m}+\frac{T_q^{\rm rem}(t)}{t^q}\right]\\
&= \left[\sum_{n=0}^{p-1}\frac{S_n(t)}{t^n}\right]\left[\sum_{m=0}^{q-1}\frac{T_m(t)}{t^m}\right]+\frac{S_p^{\rm rem}(t)}{t^p}\sum_{m=0}^{q-1}\frac{T_m(t)}{t^m}\\
&\quad+\frac{T_q^{\rm rem}(t)}{t^q}\sum_{n=0}^{p-1}\frac{S_n(t)}{t^n}+\frac{S_p^{\rm rem}(t)}{t^p}\frac{T_q^{\rm rem}(t)}{t^q}\\
&= \sum_{\ell=0}^{p-1}\frac{\sum_{n=0}^{\ell}S_n(t)T_{\ell-n}(t)}{t^\ell}+\frac{1}{t^p}\left[\sum_{\ell=p}^{p+q-2}\frac{\sum_{n=0}^{\ell}S_n(t)T_{\ell-n}(t)}{t^{\ell-p}}+\sum_{m=0}^{q-1}\frac{S_p^{\rm rem}(t)T_m(t)}{t^m}\right]\\
&\quad+\frac{1}{t^p}\left[\sum_{n=0}^{p-1}\frac{T_q^{\rm rem}(t)S_n(t)}{t^{n+q-p}}+\frac{S_p^{\rm rem}(t)T_q^{\rm rem}(t)}{t^q}\right]\\
&= \sum_{\ell=0}^{p-1}\frac{W_\ell(t)}{t^\ell}+\frac{W_p^{\rm rem}(t)}{t^p},
\end{aligned} \tag{3.104}$$

where $W_\ell(t)\equiv\sum_{n=0}^{\ell}S_n(t)T_{\ell-n}(t)$ belongs to $C_{\rm per}^{\min(r_f,r_g)-\ell}(L)$ and $W_p^{\rm rem}\in C_{\rm bdd}^{\min(r_f-p,r_g-q)}[b,\infty)$ due to the term containing the factor $S_p^{\rm rem}(t)T_q^{\rm rem}(t)$. □

Lemma 3.1.4. *For $t>0$, let*

$$f(t)\equiv\frac{1}{t^p} \tag{3.105}$$

for some real $p>0$. Also, let

$$g(t)\equiv t+\sum_{m=0}^{q-1}\frac{T_m(t)}{t^m}+\frac{T_q^{rem}(t)}{t^q}; \tag{3.106}$$

here $T_m\in C_{per}^{r-m}(L)$ and $T_q^{rem}\in C_{bdd}^{r-q}[b,\infty)$ for some integer $r\geq q$ and some real $b>0$, and we assume that $g(t)>\epsilon>0$ for $t\in[b,\infty)$. Then,

$$f\left[g(t)\right]=\frac{1}{t^p}+\frac{1}{t^{p+1}}\left[\sum_{\ell=0}^{q-1}\frac{W_\ell(t)}{t^\ell}+\frac{W_q^{rem}(t)}{t^q}\right], \tag{3.107}$$

where $W_\ell\in C_{per}^{r-\ell}(L)$ and $W_q^{rem}\in C_{bdd}^{r-q}[b,\infty)$.

Proof. Since $g(t) > \epsilon > 0$ for $t \in [b, \infty)$, it follows that $0 < f[g(t)] < \frac{1}{\epsilon^p}$ for these t. Thus,

$$\begin{aligned} f[g(t)] &= \left[t + \sum_{m=0}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q} \right]^{-p} \\ &= \frac{1}{t^p} \left\{ 1 + \frac{1}{t} \left[\sum_{m=0}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q} \right] \right\}^{-p} \end{aligned} \tag{3.108}$$

is bounded for $t \in [b, \infty)$ and can be expanded using the Taylor formula (3.99). Setting $s(x) \equiv x^{-p}$, $x \equiv 1$ and

$$a \equiv \frac{1}{t} \left[\sum_{m=0}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q} \right] \tag{3.109}$$

in that formula (with $\gamma = q + 1$), we have

$$\begin{aligned} f[g(t)] = \frac{1}{t^p} \Bigg\{ & 1 - p \frac{1}{t} \left[\sum_{m=0}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q} \right] + \frac{-p(-p-1)}{2!} \frac{1}{t^2} \left[\sum_{m=0}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q} \right]^2 + \dots \Bigg\} \\ & + \frac{1}{t^p} \frac{-p(-p-1)\dots(-p-q+1)}{q!} \frac{1}{t^q} \left[\sum_{m=0}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q} \right]^q \\ & + \frac{1}{t^p} \frac{-p(-p-1)\dots(-p-q)}{q!} \frac{1}{t^{q+1}} \left[\sum_{m=0}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q} \right]^{q+1} \\ & \cdot \int_0^1 (1-y)^q \left\{ 1 + y \frac{1}{t} \left[\sum_{m=0}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q} \right] \right\}^{-p-q-1} dy; \end{aligned} \tag{3.110}$$

applying Lemma 3.1.3 $j-1$ times for $\left[\sum_{m=0}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q} \right]^j$ and collecting terms of equal powers of $\frac{1}{t}$, we obtain

$$f[g(t)] = \frac{1}{t^p} + \frac{1}{t^{p+1}} \left[\sum_{\ell=0}^{q-1} \frac{W_\ell(t)}{t^\ell} + \frac{W_q^{\text{rem}}(t)}{t^q} \right], \tag{3.111}$$

where $W_0(t) \equiv -pT_0(t)$ belongs to $C^r_{\text{per}}(L)$, $W_1 \in C^{r-1}_{\text{per}}(L)$, etc., and $W_q^{\text{rem}} \in C^{r-q}_{\text{bdd}}[b, \infty)$ due to the term containing the factor $-pT_q^{\text{rem}}(t)$. □

Lemma 3.1.5. *For $t > 0$, let $f \in C^{r_f}_{per}(L)$, and let*

$$g(t) \equiv T_{-1} t + \sum_{m=0}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{rem}(t)}{t^q}, \tag{3.112}$$

where $r_f \geq q$, *either* $T_{-1} = 0$ *or* $T_{-1} = 1$, $T_m \in C_{per}^{r_g - m}(L)$ *and* $T_q^{rem} \in C_{bdd}^{r_g - q}[b, \infty)$ *for some integer* $r_g \geq q$ *and some real* $b > 0$. *Then,*

$$f\left[g(t)\right] = \sum_{\ell=0}^{q-1} \frac{W_\ell(t)}{t^\ell} + \frac{W_q^{rem}(t)}{t^q}, \tag{3.113}$$

where $W_\ell \in C_{per}^{\min(r_f, r_g)-\ell}(L)$ *and* $W_q^{rem} \in C_{bdd}^{\min(r_f, r_g)-q}[b, \infty)$.

Proof. We have

$$f\left[g(t)\right] = f\left[T_{-1}t + T_0(t) + \sum_{m=1}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q}\right]. \tag{3.114}$$

Using $x \equiv T_{-1}t + T_0(t)$ and

$$\begin{aligned} a &\equiv \sum_{m=1}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q} \\ &= \frac{1}{t}\left[\sum_{m=0}^{q-2} \frac{T_{m+1}(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^{q-1}}\right] \end{aligned} \tag{3.115}$$

in the Taylor formula (3.99) (with $\gamma = q$), we have

$$\begin{aligned} f\left[g(t)\right] = f\left[T_{-1}t + T_0(t)\right] &+ \frac{1}{t}\left[\sum_{m=0}^{q-2} \frac{T_{m+1}(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^{q-1}}\right] f'\left[T_{-1}t + T_0(t)\right] + \ldots \\ &+ \frac{1}{(q-1)!}\frac{1}{t^{q-1}}\left[\sum_{m=0}^{q-2} \frac{T_{m+1}(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^{q-1}}\right]^{q-1} f^{(q-1)}\left[T_{-1}t + T_0(t)\right] \\ &+ \frac{1}{(q-1)!}\frac{1}{t^q}\left[\sum_{m=0}^{q-2} \frac{T_{m+1}(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^{q-1}}\right]^{q} \\ &\cdot \int_0^1 (1-y)^{q-1} f^{(q)}\left\{T_{-1}t + T_0(t) + y\frac{1}{t}\left[\sum_{m=0}^{q-2} \frac{T_{m+1}(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^{q-1}}\right]\right\} dy. \end{aligned} \tag{3.116}$$

Applying Lemma 3.1.3 $j-1$ times for $\left[\sum_{m=0}^{q-2} \frac{T_{m+1}(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^{q-1}}\right]^j$ and collecting terms of equal powers of $\frac{1}{t}$, we therefore obtain

$$f\left[g(t)\right] = \sum_{\ell=0}^{q-1} \frac{W_\ell(t)}{t^\ell} + \frac{W_q^{\text{rem}}(t)}{t^q}, \tag{3.117}$$

where $W_0(t) \equiv f\left[T_{-1}t + T_0(t)\right]$ belongs to $C_{\text{per}}^{\min(r_f,r_g)}(L)$, $W_1(t) \equiv T_1(t)f'\left[T_{-1}t + T_0(t)\right]$ belongs to $C_{\text{per}}^{\min(r_f,r_g)-1}(L)$, etc., since $f^{(j)}\left[T_{-1}t + T_0(t)\right] \in C_{\text{per}}^{\min(r_f-j,r_g)}(L)$ and

$$f^{(q)}\left\{T_{-1}t + T_0(t) + y\left[\sum_{m=1}^{q-1}\frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q}\right]\right\} \in C_{\text{bdd}}^{\min(r_f,r_g)-q}[b,\infty).$$

□

Lemma 3.1.6. *For $t > 0$, let*

$$f(t) \equiv S_{-1}t + \sum_{n=0}^{p-1}\frac{S_n(t)}{t^n} + \frac{S_p^{rem}(t)}{t^p}, \tag{3.118}$$

where either $S_{-1} = 0$ or $S_{-1} = 1$, $S_n \in C_{per}^{r_f-n}(L)$ and $S_p^{rem} \in C_{bdd}^{r_f-p}[b,\infty)$ for some integer $r_f \geq p$ and some real $b > 0$. Also, let

$$g(t) \equiv t + \sum_{m=0}^{q-1}\frac{T_m(t)}{t^m} + \frac{T_q^{rem}(t)}{t^q}, \tag{3.119}$$

where $T_m \in C_{per}^{r_g-m}(L)$ and $T_q^{rem} \in C_{bdd}^{r_g-q}[b,\infty)$ for some integer $r_g \geq q$. Also, we assume that $g(t) > \epsilon > 0$ for $t \in [b,\infty)$. Then,

$$f\left[g(t)\right] = S_{-1}t + \sum_{\ell=0}^{\min(p,q)-1}\frac{W_\ell(t)}{t^\ell} + \frac{W_{\min(p,q)}^{rem}(t)}{t^{\min(p,q)}}, \tag{3.120}$$

where $W_\ell \in C_{per}^{\min(r_f,r_g)-\ell}(L)$ and $W_{\min(p,q)}^{rem} \in C_{bdd}^{\min(r_f-p,r_g-q)}[b,\infty)$.

Proof. We have

$$\begin{aligned} f\left[g(t)\right] = {} & S_{-1}\left[t + \sum_{m=0}^{q-1}\frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q}\right] + \sum_{n=0}^{p-1}\frac{S_n\left[t + \sum_{m=0}^{q-1}\frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q}\right]}{\left[t + \sum_{m=0}^{q-1}\frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q}\right]^n} \\ & + \frac{S_p^{\text{rem}}\left[t + \sum_{m=0}^{q-1}\frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q}\right]}{\left[t + \sum_{m=0}^{q-1}\frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q}\right]^p}. \end{aligned} \tag{3.121}$$

By Lemma 3.1.4,

$$\frac{1}{\left[t + \sum_{m=0}^{q-1}\frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q}\right]^n} \equiv \frac{1}{t^n} + \frac{1}{t^{n+1}}\left[\sum_{\ell=0}^{q-1}\frac{\hat{W}_{n_\ell}(t)}{t^\ell} + \frac{\hat{W}_{n_q}^{\text{rem}}(t)}{t^q}\right], \tag{3.122}$$

where $\hat{W}_{n_\ell} \in C_{\text{per}}^{r_g-\ell}(L)$ and $\hat{W}_{n_q}^{\text{rem}} \in C_{\text{bdd}}^{r_g-q}[b,\infty)$; by Lemma 3.1.5,

$$S_n\left[t+\sum_{m=0}^{q-1}\frac{T_m(t)}{t^m}+\frac{T_q^{\text{rem}}(t)}{t^q}\right] \equiv \sum_{\ell=0}^{q-1}\frac{\bar{W}_{n_\ell}(t)}{t^\ell}+\frac{\bar{W}_{n_q}^{\text{rem}}(t)}{t^q}, \tag{3.123}$$

where $\bar{W}_{n_\ell} \in C_{\text{per}}^{\min(r_f-n,r_g)-\ell}(L)$ and $\bar{W}_{n_q}^{\text{rem}} \in C_{\text{bdd}}^{\min(r_f-n,r_g)-q}[b,\infty)$. Thus, we have

$$\begin{aligned} f\left[g(t)\right] = {} & S_{-1}t + S_{-1}\sum_{m=0}^{q-1}\frac{T_m(t)}{t^m} + S_{-1}\frac{T_q^{\text{rem}}(t)}{t^q} + \sum_{n=0}^{p-1}\frac{1}{t^n}\left[\sum_{\ell=0}^{q-1}\frac{\bar{W}_{n_\ell}(t)}{t^\ell}+\frac{\bar{W}_{n_q}^{\text{rem}}(t)}{t^q}\right] \\ & + \sum_{n=0}^{p-1}\frac{1}{t^{n+1}}\left[\sum_{\ell=0}^{q-1}\frac{\bar{W}_{n_\ell}(t)}{t^\ell}+\frac{\bar{W}_{n_q}^{\text{rem}}(t)}{t^q}\right]\left[\sum_{\ell=0}^{q-1}\frac{\hat{W}_{n_\ell}(t)}{t^\ell}+\frac{\hat{W}_{n_q}^{\text{rem}}(t)}{t^q}\right] \\ & + S_p^{\text{rem}}\left[t+\sum_{m=0}^{q-1}\frac{T_m(t)}{t^m}+\frac{T_q^{\text{rem}}(t)}{t^q}\right]\left\{\frac{1}{t^p}+\frac{1}{t^{p+1}}\left[\sum_{\ell=0}^{q-1}\frac{\hat{W}_{p_\ell}(t)}{t^\ell}+\frac{\hat{W}_{p_q}^{\text{rem}}(t)}{t^q}\right]\right\}. \end{aligned} \tag{3.124}$$

Finally, using Lemma 3.1.3 to compute the products of sums, we conclude that

$$f\left[g(t)\right] = S_{-1}t + \sum_{\ell=0}^{\min(p,q)-1}\frac{W_\ell(t)}{t^\ell}+\frac{W_{\min(p,q)}^{\text{rem}}(t)}{t^{\min(p,q)}}, \tag{3.125}$$

where $W_0 \in C_{\text{per}}^{\min(r_f,r_g)}(L)$ due to the term $\bar{W}_{0_0}(t)$, $W_1 \in C_{\text{per}}^{\min(r_f,r_g)-1}(L)$ due to the term containing the factor $\bar{W}_{0_1}(t)$, etc., and where $W_{\min(p,q)}^{\text{rem}} \in C_{\text{bdd}}^{\min(r_f-p,r_g-q)}[b,\infty)$ due to the terms containing the factor $S_p^{\text{rem}}\left[t+\sum_{m=0}^{q-1}\frac{T_m(t)}{t^m}+\frac{T_q^{\text{rem}}(t)}{t^q}\right]$. □

Lemma 3.1.7. *For $t \in (-\infty,\infty)$, let*

$$g(t) \equiv t + T_0(t), \tag{3.126}$$

where $T_0 \in C_{per}^r(L)$ for some integer $r \geq 1$ and where we assume that $M > g'(t) > \epsilon > 0$. Then, for $t \in (-\infty,\infty)$, $g(t)$ has an inverse of the form

$$f(t) = t + S_0(t), \tag{3.127}$$

where $S_0 \in C_{per}^r(L)$.

Proof. Since $M > g'(t) > \epsilon > 0$, the function $g(t)$—which belongs to $C^r(-\infty,\infty)$—has an inverse $f \in C^r(-\infty,\infty)$ with the property $\frac{1}{\epsilon} > f'(t) > \frac{1}{M} > 0$. Defining $S_0(t) \equiv f(t) - t$,

we have

$$f(t) = t + S_0(t). \tag{3.128}$$

Thus,

$$f\left[g(t)\right] = t = t + T_0(t) + S_0\left[t + T_0(t)\right], \tag{3.129}$$

that is,

$$S_0\left[t + T_0(t)\right] = -T_0(t). \tag{3.130}$$

Since $T_0 \in C^r_{\text{per}}(L)$, this equation implies that

$$-T_0(t+L) = S_0\left[t + L + T_0(t+L)\right] = S_0\left[t + L + T_0(t)\right] \tag{3.131}$$

and

$$-T_0(t+L) = -T_0(t) = S_0\left[t + T_0(t)\right]. \tag{3.132}$$

Since $S_0 \in C^r(-\infty, \infty)$, we conclude that $S_0 \in C^r_{\text{per}}(L)$. □

Lemma 3.1.8. *For $t > 0$, let*

$$g(t) \equiv t + \sum_{m=0}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{rem}(t)}{t^q}, \tag{3.133}$$

where $T_m \in C^{r_g - m}_{per}(L)$ and $T_q^{rem} \in C^{r_g - q}_{bdd}[b, \infty)$ for some integer $r_g \geq q$ and some real $b > 0$. Also, we assume that $g(t) > \epsilon_1 > 0$ and $M > g'(t) > \epsilon_2 > 0$ for $t \in [b, \infty)$. Then, $g(t)$ has an inverse of the form

$$f(t) = t + \sum_{n=0}^{q-1} \frac{S_n(t)}{t^n} + \frac{S_q^{rem}(t)}{t^q} \tag{3.134}$$

for $t \in [g(b), \infty)$, where $S_n \in C^{r_g - n}_{per}(L)$ and $S_q^{rem} \in C^{r_g - q}_{bdd}[g(b), \infty)$.

Proof. Since $g \in C^{r_g - q}[b, \infty)$ (due to the term containing the factor $T_q^{\text{rem}}(t)$) with $M > g'(t) > \epsilon_2 > 0$, it follows that $g(t)$ has an inverse, $f \in C^{r_g - q}[g(b), \infty)$, with $\frac{1}{\epsilon_2} > f'(t) > \frac{1}{M} > 0$. Also, $g(t) \sim t + T_0(t)$ as $t \to \infty$, so based upon Lemma 3.1.7 we expect $f(t) \sim t + S_0(t)$ as $t \to \infty$ for some function $S_0 \in C^{r_g}_{\text{per}}(L)$. Thus, we seek an expansion for $f(t)$ of the form (3.134) for $t \in [g(b), \infty)$. Our method of proof is to construct sequentially the functions

$S_n(t),\ n = 0, 1, 2, \ldots q-1$ of the expansion and then show that

$$f(t) - t - \sum_{n=0}^{q-1} \frac{S_n(t)}{t^n}$$

indeed can be written as $\frac{S_q^{\text{rem}}(t)}{t^q}$ for some function $S_q^{\text{rem}} \in C_{\text{bdd}}^{r_g-q}[g(b), \infty)$.

For $t \in (-\infty, \infty)$, let

$$w(t) \equiv t + T_0(t), \tag{3.135}$$

and for $t \in [b, \infty)$ let

$$\begin{aligned} r(t) &\equiv \sum_{m=1}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q} \\ &= \frac{1}{t}\left[\sum_{m=0}^{q-2} \frac{T_{m+1}(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^{q-1}}\right] \end{aligned} \tag{3.136}$$

so that $g(t) = w(t) + r(t)$ for $t \in [b, \infty)$. We note that $w'(t) = 1 + T_0'(t)$ is L-periodic, and for any $\delta > 0$ there exists a value t_δ such that

$$\left|g'(t) - w'(t)\right| = \left|r'(t)\right| < \delta \tag{3.137}$$

for all $t > t_\delta$. We therefore may choose a δ such that $M + \delta > w'(t) > \epsilon_2 - \delta > 0$ for all $t > t_\delta$, which implies that $w(t)$ is invertible. In view of Lemma 3.1.7, we denote the inverse function as

$$w^{-1}(t) \equiv t + \bar{T}_0(t) \tag{3.138}$$

for $t \in (-\infty, \infty)$ (which includes $t \in [g(b), \infty)$), where $\bar{T}_0 \in C_{\text{per}}^{r_g}(L)$. This will be used later on in the proof.

Substituting $g(t)$ into the proposed expansion (3.134), we have (by Lemma 3.1.6)

$$\begin{aligned} f\left[g(t)\right] &= w(t) + r(t) + \sum_{n=0}^{q-1} \frac{S_n\left[w(t) + r(t)\right]}{\left[w(t) + r(t)\right]^n} + \frac{S_q^{\text{rem}}\left[w(t) + r(t)\right]}{\left[w(t) + r(t)\right]^q} \\ &= t + \sum_{\ell=0}^{q-1} \frac{W_\ell(t)}{t^\ell} + \frac{W_q^{\text{rem}}(t)}{t^q} \end{aligned} \tag{3.139}$$

for $t \in [b, \infty)$, where $W_\ell \in C_{\text{per}}^{r_g-\ell}(L)$ and $W_q^{\text{rem}} \in C_{\text{bdd}}^{r_g-q}[b, \infty)$. In particular, since $f\left[g(t)\right] =$

t, we seek functions $S_n \in C_{\text{per}}^{r_g-n}(L)$ and $S_q^{\text{rem}} \in C_{\text{bdd}}^{r_g-q}[g(b),\infty)$ such that $W_\ell(t) = 0$ for $l = 0, 1, 2, \ldots q-1$ and $W_q^{\text{rem}}(t) = 0$.

We note that

$$\frac{1}{\left[t + \sum_{m=0}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q}\right]^n} \equiv \frac{1}{t^n} + \frac{1}{t^{n+1}} \left[\sum_{\ell=0}^{q-1} \frac{\hat{W}_{n_\ell}(t)}{t^\ell} + \frac{\hat{W}_{n_q}^{\text{rem}}(t)}{t^q}\right] \tag{3.140}$$

by Lemma 3.1.4, where $\hat{W}_{n_\ell} \in C_{\text{per}}^{r_g-\ell}(L)$ and $\hat{W}_{n_q}^{\text{rem}} \in C_{\text{bdd}}^{r_g-q}[b,\infty)$. Also, as in the proof of Lemma 3.1.5, we let $x \equiv w(t)$ and $a \equiv r(t)$ in the Taylor formula (3.99) (with $\gamma = q-n$) and apply Lemma 3.1.3 $j-1$ times for $\left[\sum_{m=0}^{q-2} \frac{T_{m+1}(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^{q-1}}\right]^j$ to obtain

$$\begin{aligned}
S_n\left[w(t)+r(t)\right] \equiv\ & S_n\left[w(t)\right] + \frac{1}{t}\left[\sum_{\ell=0}^{q-2} \frac{\bar{W}_{1_{\ell+1}}(t)}{t^\ell} + \frac{\bar{W}_{1_q}^{\text{rem}}(t)}{t^{q-1}}\right] S_n'\left[w(t)\right] + \ldots \\
& + \frac{1}{(q-n-1)!}\frac{1}{t^{q-n-1}}\left[\sum_{\ell=0}^{q-2} \frac{\bar{W}_{(q-n-1)_{\ell+1}}(t)}{t^\ell} + \frac{\bar{W}_{(q-n-1)_q}^{\text{rem}}(t)}{t^{q-1}}\right] S_n^{(q-n-1)}\left[w(t)\right] \\
& + \frac{1}{(q-n-1)!}\frac{1}{t^{q-n}}\left[\sum_{\ell=0}^{q-2} \frac{\bar{W}_{(q-n)_{\ell+1}}(t)}{t^\ell} + \frac{\bar{W}_{(q-n)_q}^{\text{rem}}(t)}{t^{q-1}}\right] \\
& \cdot \int_0^1 (1-y)^{q-n-1} S_n^{(q-n)}\left[w(t)+yr(t)\right] dy,
\end{aligned} \tag{3.141}$$

where $\bar{W}_{j_{\ell+1}} \in C_{\text{per}}^{r_g-\ell-1}(L)$ and $\bar{W}_{j_q}^{\text{rem}} \in C_{\text{bdd}}^{r_g-q}[b,\infty)$. Thus,

$$\frac{S_n\left[w(t)+r(t)\right]}{\left[w(t)+r(t)\right]^n} = \frac{1}{t^n} S_n\left[w(t)\right] + \frac{1}{t^{n+1}} B_n(t), \tag{3.142}$$

where $B_n(t)$ contains derivatives of $S_n\left[w(t)\right]$ and $S_n\left[w(t)+r(t)\right]$ and is bounded for $t \in [b,\infty)$. Therefore, we construct each $S_n(t)$ as a combination of $S_0(t), S_1(t), \ldots, S_{n-1}(t)$, their derivatives and other known functions derived from $g(t)$ so as to satisfy $W_n(t) = 0$; such $S_n(t)$ necessarily belong to $C_{\text{per}}^{r_g-n}(L)$.

We apply the above expansions to (3.139) and collect terms of equal powers of $\frac{1}{t}$. First, from the $\frac{1}{t^0}$ terms we have

$$W_0(t) = T_0(t) + S_0\left[w(t)\right]. \tag{3.143}$$

Because $w(t)$ has the inverse $w^{-1}(t)$, we set

$$S_0(t) \equiv -T_0\left[w^{-1}(t)\right] = -T_0\left[t + \bar{T}_0(t)\right] \tag{3.144}$$

so that

$$W_0(t) = T_0(t) + S_0\left[w(t)\right] = 0. \tag{3.145}$$

Similarly, given the $\frac{1}{t}$ terms, we set

$$S_1(t) \equiv -\hat{W}_{0_0}\left[w^{-1}(t)\right] S_0(t) - \bar{W}_{1_1}\left[w^{-1}(t)\right] S_0'(t) - T_1\left[w^{-1}(t)\right] \tag{3.146}$$

so that

$$W_1(t) = T_1(t) + \hat{W}_{0_0}(t) S_0\left[w(t)\right] + \bar{W}_{1_1}(t) S_0'\left[w(t)\right] + S_1\left[w(t)\right] = 0. \tag{3.147}$$

Continuing this process up through the equation $W_{q-1}(t) = 0$, we construct functions $S_n(t),\ n = 0, 1, 2, \ldots q-1$ such that $S_n \in C^{r_g - n}_{\text{per}}(L)$.

Given these choices for $S_n(t),\ n = 0, 1, 2, \ldots q-1$, (3.139) can be re-expressed as

$$\frac{W_q^{\text{rem}}(t)}{t^q} = \frac{S_q^{\text{rem}}\left[w(t) + r(t)\right]}{\left[t + \sum_{m=0}^{q-1} \frac{T_m(t)}{t^m} + \frac{T_q^{\text{rem}}(t)}{t^q}\right]^q} + \frac{1}{t^q}\bar{B}_q(t), \tag{3.148}$$

where $\bar{B}_q \in C^{r_g - q}_{\text{bdd}}[b, \infty)$ is a combination of $S_0(t), S_1(t), \ldots, S_{q-1}(t)$, their derivatives and other known functions derived from $g(t)$. We now show that there, in fact, exists a function $S_q^{\text{rem}} \in C^{r_g - q}_{\text{bdd}}[g(b), \infty)$ such that

$$W_q^{\text{rem}}(t) = \frac{S_q^{\text{rem}}\left[w(t) + r(t)\right]}{\left[1 + \sum_{m=0}^{q-1} \frac{T_m(t)}{t^{m+1}} + \frac{T_q^{\text{rem}}(t)}{t^{q+1}}\right]^q} + \bar{B}_q(t) = 0. \tag{3.149}$$

Since

$$\left[1 + \sum_{m=0}^{q-1} \frac{T_m(t)}{t^{m+1}} + \frac{T_q^{\text{rem}}(t)}{t^{q+1}}\right]^q = \frac{[g(t)]^q}{t^q} \tag{3.150}$$

is $C^{r_g - q}_{\text{bdd}}[b, \infty)$, the function

$$-\left[1 + \sum_{m=0}^{q-1} \frac{T_m(t)}{t^{m+1}} + \frac{T_q^{\text{rem}}(t)}{t^{q+1}}\right]^q \bar{B}_q(t)$$

also belongs to $C^{r_g-q}_{\text{bdd}}[b,\infty)$, and since $f \in C^{r_g}[b,\infty)$, it follows that

$$-\left[1+\sum_{m=0}^{q-1}\frac{T_m\left[f(t)\right]}{t^{m+1}}+\frac{T_q^{\text{rem}}\left[f(t)\right]}{t^{q+1}}\right]^q \bar{B}_q\left[f(t)\right]$$

is $C^{r_g-q}_{\text{bdd}}[b,\infty)$. Thus, there exists a function $S_q^{\text{rem}} \in C^{r_g-q}_{\text{bdd}}[g(b),\infty)$, defined implicitly as

$$S_q^{\text{rem}}(t) \equiv -\left[1+\sum_{m=0}^{q-1}\frac{T_m\left[f(t)\right]}{t^{m+1}}+\frac{T_q^{\text{rem}}\left[f(t)\right]}{t^{q+1}}\right]^q \bar{B}_q\left[f(t)\right], \tag{3.151}$$

such that (3.149) holds. We conclude that $f(t)$ indeed can be expanded according to the formula (3.134). □

3.1.3.3 Proof of Theorem 3.1.9

Having established the lemmas of the previous section, we now prove Theorem 3.1.9.

Proof. We consider each of the three parts of the theorem individually. After an in-depth discussion for Part 1, we briefly describe how the proof of Part 2 follows analogously. We then show how Part 3 proceeds straightforwardly from these results.

Part 1: For this part, we give all of the details of the proof for the integral $I_+^2(x,k,c,A)$ (3.94), which involves estimating the size of the integral $I_n(x,k,c,A)$ (3.98) for every $n \in \mathbb{Z}$. At the end, we outline how the proof for the integral $I_+^1(x,k,c,A)$ follows similarly.

The integrand in (3.98). We start by using the lemmas of Section 3.1.3.2 to show that the integrand in (3.98) can be expanded as a sum of smoothly decaying terms containing periodic functions of ξ plus an additional smoothly decaying remainder term.

First, applying the Taylor formula (3.99) to $\tilde{\phi}_+(x,t)$, we have

$$\begin{aligned}
\xi &= \tilde{\phi}_+(x,t) \\
&= \frac{1}{1+\sin(\theta)+c_1}\left[t\sqrt{1+\left(\frac{h}{L}\right)^2\left\{\frac{\Delta}{t}\right\}^2}+(\sin(\theta),-\cos(\theta))\cdot\left(t,\left(\frac{h}{L}\right)\Delta\right)+c_1 t\right] \\
&= t+\frac{1}{1+\sin(\theta)+c_1}\left[-\cos(\theta)\left(\frac{h}{L}\right)\Delta+t\frac{1}{2}\left(\frac{h}{L}\right)^2\frac{\Delta^2}{t^2}+t\frac{\left(\frac{1}{2}\right)\left(-\frac{1}{2}\right)}{2!}\left(\frac{h}{L}\right)^4\frac{\Delta^4}{t^4}+\ldots\right] \\
&\quad+\frac{t}{1+\sin(\theta)+c_1}\frac{\left(\frac{1}{2}\right)\left(-\frac{1}{2}\right)\ldots\left(\frac{1}{2}-q+2\right)}{(q-1)!}\left(\frac{h}{L}\right)^{2q-2}\frac{\Delta^{2q-2}}{t^{2q-2}} \\
&\quad+\frac{t}{1+\sin(\theta)+c_1}\frac{\left(\frac{1}{2}\right)\left(-\frac{1}{2}\right)\ldots\left(\frac{1}{2}-q+1\right)}{(q-1)!}\left(\frac{h}{L}\right)^{2q}\frac{\Delta^{2q}}{t^{2q}}\int_0^1(1-y)^{q-1}\left(1+y\left(\frac{h}{L}\right)^2\frac{\Delta^2}{t^2}\right)^{\frac{1}{2}-q}dy \\
&= t+\sum_{m=0}^{2q-2}\frac{Q_m\left(\frac{2\pi x}{L},2\pi t\right)}{t^m}+\frac{Q_{2q-1}^{\mathrm{rem}}\left(\frac{2\pi x}{L},2\pi t\right)}{t^{2q-1}}
\end{aligned} \tag{3.152}$$

for some arbitrarily large integer $q\geq 1$, where

$$\begin{aligned}
\Delta &\equiv \Delta\left(\frac{2\pi x}{L},2\pi t\right) \\
&\equiv \frac{\bar{f}\left[2\pi\left(\frac{x}{L}+t\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)}{2}.
\end{aligned} \tag{3.153}$$

Here each $Q_m\left(\frac{2\pi x}{L},2\pi t\right)$ is $C^\infty_{\mathrm{per}}(L)$ in x and $C^\infty_{\mathrm{per}}(1)$ in t, while $Q^{\mathrm{rem}}_{2q-1}\left(\frac{2\pi x}{L},2\pi t\right)$ is $C^\infty_{\mathrm{per}}(L)$ in x and $C^\infty_{\mathrm{bdd}}\left[\frac{cA}{L},\infty\right)$ in t, since all of these functions are directly determined in terms of the $C^\infty_{\mathrm{per}}(L)$ grating profile $f(x)$ using the above equation. In particular, $Q_2=Q_4=\ldots=Q_{2q-2}=0$. Also, $M>\frac{\partial\tilde{\phi}_+}{\partial t}>\epsilon_2>0$ for $t\in(0,\infty)$ with $\lim_{t\to 0}\tilde{\phi}_+(x,t)=0$, which implies that $M>\frac{\partial\tilde{\phi}_+}{\partial t}>\epsilon_2>0$ and $\tilde{\phi}_+(x,t)>\epsilon_1>0$ for $t\in\left[\frac{cA}{L},\infty\right)$.

Remark 3.1.5. *For every integer $p\geq 1$, we choose q to be sufficiently large relative to p so as to allow us to compute the estimate (3.77) for $I^2_+(x,k,c,A)$.*

Thus, by Lemma 3.1.8, we may write

$$\begin{aligned}
t &= \tilde{\phi}_+^{-1}(x,\xi) \\
&= \xi+\sum_{m=0}^{2q-2}\frac{R_m\left(\frac{2\pi x}{L},2\pi\xi\right)}{\xi^m}+\frac{R^{\mathrm{rem}}_{2q-1}\left(\frac{2\pi x}{L},2\pi\xi\right)}{\xi^{2q-1}}
\end{aligned} \tag{3.154}$$

for $\xi \in \left[\tilde{\phi}_+\left(x, \frac{cA}{L}\right), \infty\right)$, where $R_m\left(\frac{2\pi x}{L}, 2\pi\xi\right) \in C^\infty_{\text{per}}(1)$ in ξ, $R^{\text{rem}}_{2q-1}\left(\frac{2\pi x}{L}, 2\pi\xi\right) \in C^\infty_{\text{bdd}}\left[\tilde{\phi}_+\left(x, \frac{cA}{L}\right), \infty\right)$ in ξ and $\frac{1}{\epsilon_2} > \frac{\partial \tilde{\phi}_+^{-1}}{\partial \xi} > \frac{1}{M} > 0$. We note that because $Q_m\left(\frac{2\pi x}{L}, 2\pi t\right)$ and $Q^{\text{rem}}_{2q-1}\left(\frac{2\pi x}{L}, 2\pi t\right)$ are $C^\infty_{\text{per}}(L)$ in x, $R_m\left(\frac{2\pi x}{L}, 2\pi\xi\right)$ and $R^{\text{rem}}_{2q-1}\left(\frac{2\pi x}{L}, 2\pi\xi\right)$ are $C^\infty_{\text{per}}(L)$ in x as well. This key formula for $\tilde{\phi}_+^{-1}(x, \xi)$, which contains terms that are periodic functions of ξ divided by integer powers of ξ, allows us to expand the integrand of $I_n(x, k, c, A)$ in a similar manner—ultimately leading to an estimation via integration by parts of the size in $\frac{A}{L}$ of the integral as $A \to \infty$.

By (3.87), we have

$$\bar{g}_+\left(x, \tilde{\phi}_+^{-1}\right) = \frac{i}{2}\frac{h}{2L}\frac{1}{\tilde{\phi}_+^{-1}}\frac{\frac{\bar{f}\left[2\pi\left(\frac{x}{L}+\tilde{\phi}_+^{-1}\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)}{\tilde{\phi}_+^{-1}} - 2\pi\bar{f}'\left(\frac{2\pi x}{L}\right)}{1+\left(\frac{h}{2L}\right)^2\left\{\frac{\bar{f}\left[2\pi\left(\frac{x}{L}+\tilde{\phi}_+^{-1}\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)}{\tilde{\phi}_+^{-1}}\right\}^2}, \tag{3.155}$$

and

$$\begin{aligned}\frac{h\left[kL\bar{u}_+\left(x, \tilde{\phi}_+^{-1}\right)\right]}{e^{ikL\bar{u}_+\left(x, \tilde{\phi}_+^{-1}\right)}} &= \frac{kL\bar{u}_+ H_1^1\left(kL\bar{u}_+\right)}{e^{ikL\bar{u}_+}} \\ &\sim \left(\frac{2}{\pi}\right)^{\frac{1}{2}} e^{-i\frac{3\pi}{4}}\left(kL\bar{u}_+\right)^{\frac{1}{2}}\sum_{m=0}^{p-1}\frac{(-1)^m\Gamma\left(\frac{3}{2}+m\right)}{m!\Gamma\left(\frac{3}{2}-m\right)\left(2ikL\bar{u}_+\right)^m} \\ &\quad + \mathcal{O}\left(\left(kL\bar{u}_+\right)^{\frac{1}{2}-p}\right)\end{aligned} \tag{3.156}$$

as $\xi \to \infty$ by (2.62) and (2.86) (for any integer $p \geq 1$); here, in view of (3.88),

$$\bar{u}_+\left(x, \tilde{\phi}_+^{-1}\right) = \tilde{\phi}_+^{-1}\sqrt{1+\left(\frac{h}{2L}\right)^2\left\{\frac{\bar{f}\left[2\pi\left(\frac{x}{L}+\tilde{\phi}_+^{-1}\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)}{\tilde{\phi}_+^{-1}}\right\}^2}. \tag{3.157}$$

Thus, it follows that

$$\begin{aligned}\bar{g}_+\left(x,\tilde{\phi}_+^{-1}\right)\frac{h\left[kL\bar{u}_+\left(x,\tilde{\phi}_+^{-1}\right)\right]}{e^{ikL\bar{u}_+\left(x,\tilde{\phi}_+^{-1}\right)}} &\sim \frac{h}{L}\sqrt{\frac{kL}{\tilde{\phi}_+^{-1}}}\frac{i}{4}\frac{\frac{\bar{f}\left[2\pi\left(\frac{x}{L}+\tilde{\phi}_+^{-1}\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)}{\tilde{\phi}_+^{-1}}-2\pi\bar{f}'\left(\frac{2\pi x}{L}\right)}{\left(1+\left(\frac{h}{2L}\right)^2\left\{\frac{\bar{f}\left[2\pi\left(\frac{x}{L}+\tilde{\phi}_+^{-1}\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)}{\tilde{\phi}_+^{-1}}\right\}^2\right)^{\frac{3}{4}}}\\ &\quad\times\left(\frac{2}{\pi}\right)^{\frac{1}{2}}e^{-i\frac{3\pi}{4}}\sum_{m=0}^{p-1}\frac{(-1)^m\Gamma\left(\frac{3}{2}+m\right)}{m!\Gamma\left(\frac{3}{2}-m\right)(2i)^m}\left(kL\bar{u}_+\right)^{-m}\\ &\quad+\frac{h}{L}\sqrt{\frac{kL}{\tilde{\phi}_+^{-1}}}\times\mathcal{O}\left(\left(kL\bar{u}_+\right)^{-p}\right)\end{aligned} \tag{3.158}$$

as $\xi\to\infty$. Keeping in mind (3.154), we also have

$$e^{i(2\pi n-kLc_1)\tilde{\phi}_+^{-1}}=e^{i(2\pi n-kLc_1)\xi}e^{i(2\pi n-kLc_1)\left[\sum_{m=0}^{2q-2}\frac{R_m\left(\frac{2\pi x}{L},2\pi\xi\right)}{\xi^n}+\frac{R_{2q-1}^{\mathrm{rem}}\left(\frac{2\pi x}{L},2\pi\xi\right)}{\xi^{2q-1}}\right]}; \tag{3.159}$$

the second term on the right-hand side can be expanded using Lemma 3.1.5. The last function in the integrand of (3.98) is

$$\begin{aligned}\frac{\partial\tilde{\phi}_+^{-1}}{\partial\xi}&=\frac{\partial}{\partial\xi}\left[\xi+\sum_{m=0}^{2q-2}\frac{R_m\left(\frac{2\pi x}{L},2\pi\xi\right)}{\xi^m}+\frac{R_{2q-1}^{\mathrm{rem}}\left(\frac{2\pi x}{L},2\pi\xi\right)}{\xi^{2q-1}}\right]\\ &=\sum_{m=0}^{2q-2}\frac{\bar{R}_m\left(\frac{2\pi x}{L},2\pi\xi\right)}{\xi^m}+\frac{\bar{R}_{2q-1}^{\mathrm{rem}}\left(\frac{2\pi x}{L},2\pi\xi\right)}{\xi^{2q-1}},\end{aligned} \tag{3.160}$$

where each $\bar{R}_m\left(\frac{2\pi x}{L},2\pi\xi\right)$ is $C_{\mathrm{per}}^\infty(L)$ in x and $C_{\mathrm{per}}^\infty(1)$ in ξ and where $\bar{R}_{2q-1}^{\mathrm{rem}}\left(\frac{2\pi x}{L},2\pi\xi\right)$ is $C_{\mathrm{per}}^\infty(L)$ in x and $C_{\mathrm{bdd}}^\infty\left[\tilde{\phi}_+\left(x,\frac{cA}{L}\right),\infty\right)$ in ξ. Therefore, choosing q to be sufficiently large relative to p (Remark 3.1.5), using Lemmas (3.1.3)–(3.1.5) and noting that the expansion for $\left(\tilde{\phi}_+^{-1}\right)^{-m-\frac{1}{2}}$ has $\xi^{-m-\frac{1}{2}}$ as its leading order term, we may write

$$\begin{aligned}&\bar{g}_+\left(x,\tilde{\phi}_+^{-1}\right)\frac{h\left[kL\bar{u}_+\left(x,\tilde{\phi}_+^{-1}\right)\right]}{e^{ikL\bar{u}_+\left(x,\tilde{\phi}_+^{-1}\right)}}e^{i(2\pi n-kLc_1)\tilde{\phi}_+^{-1}}\frac{\partial\tilde{\phi}_+^{-1}}{\partial\xi}\\ &\quad=e^{i(2\pi n-kLc_1)\xi}\left[\sum_{\ell=0}^{p-1}\frac{V_{n,\ell}\left(kL,\frac{2\pi x}{L},2\pi\xi\right)}{\xi^{\ell+\frac{1}{2}}}+\frac{V_{n,p}^{\mathrm{rem}}\left(kL,\frac{2\pi x}{L},2\pi\xi\right)}{\xi^{p+\frac{1}{2}}}\right]\end{aligned} \tag{3.161}$$

for $\xi\in\left[\tilde{\phi}_+\left(x,\frac{cA}{L}\right),\infty\right)$, where each $V_{n,\ell}\left(kL,\frac{2\pi x}{L},2\pi\xi\right)$ is $C_{\mathrm{per}}^\infty(L)$ in x and $C_{\mathrm{per}}^\infty(1)$ in ξ

and each $V_{n,p}^{\mathrm{rem}}\left(kL, \frac{2\pi x}{L}, 2\pi\xi\right)$ is $C_{\mathrm{per}}^{\infty}(L)$ in x and $C_{\mathrm{bdd}}^{\infty}\left[\tilde{\phi}_+\left(x, \frac{cA}{L}\right), \infty\right)$ in ξ, and we re-express (3.98) as

$$\begin{aligned} I_n(x,k,c,A) = d_n e^{in\frac{2\pi x}{L}} \int_{\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty} \bar{P}_2\left(c, \frac{L\tilde{\phi}_+^{-1}}{cA}\right) \sum_{\ell=0}^{p-1} \frac{V_{n,\ell}\left(kL, \frac{2\pi x}{L}, 2\pi\xi\right)}{\xi^{\ell+\frac{1}{2}}} e^{i\{kL[1+\sin(\theta)]+2\pi n\}\xi}\, d\xi \\ + d_n e^{in\frac{2\pi x}{L}} \int_{\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty} \bar{P}_2\left(c, \frac{L\tilde{\phi}_+^{-1}}{cA}\right) \frac{V_{n,p}^{\mathrm{rem}}\left(kL, \frac{2\pi x}{L}, 2\pi\xi\right)}{\xi^{p+\frac{1}{2}}} e^{i\{kL[1+\sin(\theta)]+2\pi n\}\xi}\, d\xi. \end{aligned} \tag{3.162}$$

Fourier expansion. Using the Fourier expansion

$$V_{n,\ell}\left(kL, \frac{2\pi x}{L}, 2\pi\xi\right) = \sum_{m=-\infty}^{\infty} v_{n,\ell,m}\left(kL, \frac{2\pi x}{L}\right) e^{i2\pi m\xi}, \tag{3.163}$$

we define

$$\begin{aligned} &I_{n,\ell,m}(x,k,c,A) \\ &\quad \equiv d_n e^{in\frac{2\pi x}{L}} v_{n,\ell,m}\left(kL, \frac{2\pi x}{L}\right) \int_{\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty} \frac{\bar{P}_2\left(c, \frac{L\tilde{\phi}_+^{-1}}{cA}\right)}{\xi^{\ell+\frac{1}{2}}} e^{i\{kL[1+\sin(\theta)]+2\pi(n+m)\}\xi}\, d\xi \end{aligned} \tag{3.164}$$

and

$$\begin{aligned} &I_n^{\mathrm{rem}}(x,k,c,A) \\ &\quad \equiv d_n e^{in\frac{2\pi x}{L}} \int_{\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty} \bar{P}_2\left(c, \frac{L\tilde{\phi}_+^{-1}}{cA}\right) \frac{V_{n,p}^{\mathrm{rem}}\left(kL, \frac{2\pi x}{L}, 2\pi\xi\right)}{\xi^{p+\frac{1}{2}}} e^{i\{kL[1+\sin(\theta)]+2\pi n\}\xi}\, d\xi. \end{aligned} \tag{3.165}$$

Analogous to the Fourier series (3.95), the series (3.163) is uniformly convergent for $(x,\xi) \in [0,L]\times\left[\tilde{\phi}_+\left(x, \frac{cA}{L}\right), \infty\right)$ with coefficients $v_{n,\ell,m}$ that (in magnitude) decrease super-algebraically in m as $m \to \pm\infty$ (for each n,ℓ) due to the smoothness of the grating profile. Therefore, we have

$$\begin{aligned} I_+^2(x,k,c,A) &= \sum_{n=-\infty}^{\infty} I_n(x,k,c,A) \\ &= \sum_{n=-\infty}^{\infty} \left\{ \sum_{\ell=0}^{p-1} \sum_{m=-\infty}^{\infty} I_{n,\ell,m}(x,k,c,A) + I_n^{\mathrm{rem}}(x,k,c,A) \right\}. \end{aligned} \tag{3.166}$$

For each ℓ, the functions $V_{n,\ell}\left(kL, \frac{2\pi x}{L}, 2\pi\xi\right)$ contain the factor n^s for some integer $s \in \{0, 1, 2, \ldots, 2q-1\}$. This is because the expansion of (3.159) that is generated by using Lemma 3.1.5 includes derivatives with respect to z of $e^{i(2\pi n - kLc_1)z}$ up to order $2q-1$, as can be deduced from the formula (3.116) in the proof of Lemma 3.1.5 together with the sum $\sum_{m=0}^{2q-2} \frac{R_m\left(\frac{2\pi x}{L}, 2\pi\xi\right)}{\xi^n} + \frac{R^{\text{rem}}_{2q-1}\left(\frac{2\pi x}{L}, 2\pi\xi\right)}{\xi^{2q-1}}$ in (3.159). Thus, the coefficients $v_{n,\ell,m}$ have the property

$$\left| v_{n,\ell,m}\left(kL, \frac{2\pi x}{L}\right)\right| \leq C_{\ell,q,r}(x,k)|n|^{2q-1}|m|^{-r} \tag{3.167}$$

for all $n, m \in \mathbb{Z}$ for every integer $r \geq 1$ and certain positive real functions $C_{\ell,q,r}(x,k)$ that are independent of n, m. Similarly, the functions $V^{\text{rem}}_{n,p}\left(kL, \frac{2\pi x}{L}, 2\pi\xi\right)$—which are $C^\infty_{\text{bdd}}\left[\tilde{\phi}_+\left(x, \frac{cA}{L}\right), \infty\right)$ in ξ—satisfy

$$\left| V^{\text{rem}}_{n,p}\left(kL, \frac{2\pi x}{L}, 2\pi\xi\right)\right| \leq C_p(x,k)|n|^{2q-1} \tag{3.168}$$

for all $n \in \mathbb{Z}$ for some positive real function $C_p(x,k)$ that is independent of n. Since the coefficients d_n (in magnitude) decrease super-algebraically in n as $n \to \pm\infty$ (3.96), however, the series in (3.166) are absolutely and uniformly convergent (as we will demonstrate explicitly later). So, estimates for $|I_{n,\ell,m}(x,k,c,A)|$ and $|I^{\text{rem}}_n(x,k,c,A)|$ can be turned into estimates for $|I_n(x,k,c,A)|$ and thus $\left|I^2_+(x,k,c,A)\right|$.

Change of variables. To analyze $I_{n,\ell,m}(x,k,c,A)$ and $I^{\text{rem}}_n(x,k,c,A)$, we introduce a change of variables additional to the formula (3.91), namely $\bar{\xi} \equiv \frac{L}{cA}\xi$, and we write

$$\begin{aligned} I_{n,\ell,m}(x,k,c,A) = d_n e^{in\frac{2\pi x}{L}} v_{n,\ell,m}\left(kL, \frac{2\pi x}{L}\right) \frac{1}{\left(\frac{cA}{L}\right)^{\ell-\frac{1}{2}}} \\ \times \int_{\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty} \frac{\bar{P}_2\left[c, \frac{L}{cA}\tilde{\phi}_+^{-1}\left(x, \frac{cA}{L}\bar{\xi}\right)\right]}{\bar{\xi}^{\ell+\frac{1}{2}}} e^{i\{kL[1+\sin(\theta)]+2\pi(n+m)\}\frac{cA}{L}\bar{\xi}}\, d\bar{\xi} \end{aligned} \tag{3.169}$$

and

$$\begin{aligned} I^{\text{rem}}_n(x,k,c,A) = d_n e^{in\frac{2\pi x}{L}} \frac{1}{\left(\frac{cA}{L}\right)^{p-\frac{1}{2}}} \int_{\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty} \bar{P}_2\left[c, \frac{L}{cA}\tilde{\phi}_+^{-1}\left(x, \frac{cA}{L}\bar{\xi}\right)\right] \\ \times \frac{V^{\text{rem}}_{n,p}\left(kL, \frac{2\pi x}{L}, 2\pi\frac{cA}{L}\bar{\xi}\right)}{\bar{\xi}^{p+\frac{1}{2}}} e^{i\{kL[1+\sin(\theta)]+2\pi n\}\frac{cA}{L}\bar{\xi}}\, d\bar{\xi}. \end{aligned} \tag{3.170}$$

These forms are useful for large values of $\frac{A}{L}$: the lower limit of integration for these integrals is

$$\begin{aligned}\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right) &= \frac{L}{cA}\left[\frac{cA}{L}+\sum_{m=0}^{2q-2}\frac{Q_m\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\right)}{\left(\frac{cA}{L}\right)^m}+\frac{Q^{\rm rem}_{2q-1}\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\right)}{\left(\frac{cA}{L}\right)^{2q-1}}\right]\\ &= 1+\sum_{m=0}^{2q-2}\frac{Q_m\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\right)}{\left(\frac{cA}{L}\right)^{m+1}}+\frac{Q^{\rm rem}_{2q-1}\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\right)}{\left(\frac{cA}{L}\right)^{2q}},\end{aligned} \tag{3.171}$$

which is $\mathcal{O}(1)$ in $\frac{A}{L}$ as $A\to\infty$, and for

$$\bar{P}_2\left[c,\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)\right] = 1-S\left(\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right),1,\frac{1}{c}\right) \tag{3.172}$$

we have

$$\begin{aligned}\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right) &= \frac{L}{cA}\left[\frac{cA}{L}\bar{\xi}+\sum_{m=0}^{2q-2}\frac{R_m\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\left(\frac{cA}{L}\bar{\xi}\right)^m}+\frac{R^{\rm rem}_{2q-1}\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\left(\frac{cA}{L}\bar{\xi}\right)^{2q-1}}\right]\\ &= \bar{\xi}+\sum_{m=0}^{2q-2}\frac{R_m\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\frac{cA}{L}\left(\frac{cA}{L}\bar{\xi}\right)^m}+\frac{R^{\rm rem}_{2q-1}\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\frac{cA}{L}\left(\frac{cA}{L}\bar{\xi}\right)^{2q-1}},\end{aligned} \tag{3.173}$$

which also is $\mathcal{O}(1)$ in $\frac{A}{L}$ as $A\to\infty$.

Bound on $I_n^{\rm rem}$. Considering $I_n^{\rm rem}(x,k,c,A)$ first, we note that the integral in (3.170) is absolutely convergent because $\bar{P}_2\left[c,\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)\right]V^{\rm rem}_{n,p}\left(kL,\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)$ is $C^\infty_{\rm bdd}\left[\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right),\infty\right)$ in $\bar{\xi}$ and the function $\frac{1}{\bar{\xi}^{p+\frac{1}{2}}}$ decays sufficiently rapidly for $p\geq 1$. Since $\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right)$ and $\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)$ are $\mathcal{O}(1)$ in $\frac{A}{L}$ as $A\to\infty$, we therefore have the estimate

$$\begin{aligned}|I_n^{\rm rem}(x,k,c,A)| &= \frac{|d_n|}{\left(\frac{cA}{L}\right)^{p-\frac{1}{2}}}\left|\int_{\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^\infty \bar{P}_2\left[c,\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)\right]\right.\\ &\qquad\left.\times\frac{V^{\rm rem}_{n,p}\left(kL,\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\bar{\xi}^{p+\frac{1}{2}}}e^{i\{kL[1+\sin(\theta)]+2\pi n\}\frac{cA}{L}\bar{\xi}}\,d\bar{\xi}\right|\\ &\leq \frac{|d_n|}{\left(\frac{cA}{L}\right)^{p-\frac{1}{2}}}\int_{\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^\infty\left|\bar{P}_2\left[c,\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)\right]\frac{V^{\rm rem}_{n,p}\left(kL,\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\bar{\xi}^{p+\frac{1}{2}}}\right|d\bar{\xi}\\ &\leq |d_n|M^{\rm rem}_{n,p}(x,k,c)\left(\frac{A}{L}\right)^{-p+\frac{1}{2}}\end{aligned} \tag{3.174}$$

as $A \to \infty$, where $M_{n,p}^{\mathrm{rem}}(x,k,c)$ is some positive real function of x,k,c that is independent of A; due to the bound (3.168) for $V_{n,p}^{\mathrm{rem}}\left(kL, \frac{2\pi x}{L}, 2\pi\xi\right)$, the functions $M_{n,p}^{\mathrm{rem}}(x,k,c)$ satisfy

$$M_{n,p}^{\mathrm{rem}}(x,k,c) \leq \bar{M}_p(x,k,c)|n|^{2q-1} \tag{3.175}$$

for all $n \in \mathbb{Z}$ for some positive real function $\bar{M}_p(x,k,c)$ that is independent of n.

Bound on $I_{n,\ell,m}$. We now show that $I_{n,\ell,m}(x,k,c,A)$ also is $\mathcal{O}\left(\left(\frac{A}{L}\right)^{-p+\frac{1}{2}}\right)$ as $A \to \infty$. This is done by applying integration by parts $p-\ell$ times ($1 \leq p-\ell \leq p$) to the integral in (3.169); in particular, we will integrate the factors that are periodic in $\bar{\xi}$ and differentiate the factors which contain the POU function $\bar{P}_2$ and an inverse power of $\bar{\xi}$, and after each integration by parts iteration we will expand the resulting integrand terms as necessary so that the process can be repeated. This ensures that at each iteration we gain an additional $\left(\frac{A}{L}\right)^{-1}$ factor in our estimation of the size of the integral until we reach the desired result.

In the first application of integration by parts, we differentiate $\frac{\bar{P}_2\left[c, \frac{L}{cA}\tilde{\phi}_+^{-1}\left(x, \frac{cA}{L}\bar{\xi}\right)\right]}{\bar{\xi}^{\ell+\frac{1}{2}}}$ and integrate $e^{i\{kL[1+\sin(\theta)]+2\pi(n+m)\}\frac{cA}{L}\bar{\xi}}$. The integration is straightforward: here we invoke the assumption that $|kL\left[1+\sin(\theta)\right]+2\pi n| > \eta > 0$ for every $n \in \mathbb{Z}$ so that the antiderivative is guaranteed to exist for every $n,m \in \mathbb{Z}$. The differentiation is somewhat more involved, however, due to the presence of $\bar{P}_2\left[c, \frac{L}{cA}\tilde{\phi}_+^{-1}\left(x, \frac{cA}{L}\bar{\xi}\right)\right]$. Differentiating (3.172), we have

$$\frac{\partial \bar{P}_2\left[c, \frac{L}{cA}\tilde{\phi}_+^{-1}\left(x, \frac{cA}{L}\bar{\xi}\right)\right]}{\partial \bar{\xi}} = -\left.\frac{\partial S\left(z,1,\frac{1}{c}\right)}{\partial z}\right|_{z=\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)} \frac{\partial}{\partial \bar{\xi}}\left[\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x, \frac{cA}{L}\bar{\xi}\right)\right]. \tag{3.176}$$

Now,

$$\frac{\partial S\left(z,1,\frac{1}{c}\right)}{\partial z}\begin{cases} 0 & , \ 0 < z \leq 1 \\ \frac{\partial \exp\left(\frac{2e^{-1/u}}{u-1}\right)}{\partial u}\frac{1}{\frac{1}{c}-1} & , \ 1 < z < \frac{1}{c}, \quad u = \frac{z-1}{\frac{1}{c}-1} \\ 0 & , \ z \geq \frac{1}{c}; \end{cases} \tag{3.177}$$

for $z = \frac{L}{cA}\tilde{\phi}_+^{-1}\left(x, \frac{cA}{L}\bar{\xi}\right)$, this factor is bounded in $1 < z < \frac{1}{c}$ (the finite interval in which the derivative is nonzero) because $c > 0$ is a fixed constant that is strictly less than 1, and—as

shown by (3.173)—it is $\mathcal{O}(1)$ in $\frac{A}{L}$ as $A \to \infty$. Since

$$\begin{aligned}\frac{\partial}{\partial\bar{\xi}}\left[\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)\right] &= \frac{\partial}{\partial\xi}\left[\tilde{\phi}_+^{-1}(x,\xi)\right]\Big|_{\xi=\frac{cA}{L}\bar{\xi}} \\ &= \sum_{m=0}^{2q-2}\frac{\bar{R}_m\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\left(\frac{cA}{L}\bar{\xi}\right)^m} + \frac{\bar{R}^{\text{rem}}_{2q-1}\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\left(\frac{cA}{L}\bar{\xi}\right)^{2q-1}},\end{aligned} \tag{3.178}$$

where each $\bar{R}_m\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)$ is $C^\infty_{\text{per}}(L)$ in x and $C^\infty_{\text{per}}\left(\frac{L}{cA}\right)$ in $\bar{\xi}$ and $\bar{R}^{\text{rem}}_{2q-1}\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)$ is $C^\infty_{\text{per}}(L)$ in x and $C^\infty_{\text{bdd}}\left[\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right),\infty\right)$ in $\bar{\xi}$, we conclude that $\frac{\partial\bar{P}_2\left[c,\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)\right]}{\partial\bar{\xi}}$ is $\mathcal{O}(1)$ in $\frac{A}{L}$ as $A\to\infty$.

We note that the boundary term which is generated by the first integration by parts iteration is equal to 0, since $\bar{P}_2\left[c,\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)\right]$ and its derivative is 0 at $\bar{\xi}=\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right)$ while $\frac{1}{\bar{\xi}^{\ell+\frac{1}{2}}}$ and its derivative decays to 0 as $\bar{\xi}\to\infty$. Thus, integrating by parts once, we reformulate $I_{n,\ell,m}(x,k,c,A)$ as

$$\begin{aligned}I_{n,\ell,m}(x,k,c,A) = {}& \frac{-d_n e^{in\frac{2\pi x}{L}}v_{n,\ell,m}\left(kL,\frac{2\pi x}{L}\right)}{i\left\{kL\left[1+\sin(\theta)\right]+2\pi(n+m)\right\}}\frac{1}{\left(\frac{cA}{L}\right)^{\ell+\frac{1}{2}}} \\ &\times\int_{\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty}\left\{\frac{\left(-\ell-\frac{1}{2}\right)\bar{P}_2\left[c,\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)\right]}{\bar{\xi}^{\ell+\frac{3}{2}}}\right. \\ &\left.+\frac{-\frac{\partial S\left(z,1,\frac{1}{c}\right)}{\partial z}\Big|_{z=\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)}\left[\sum_{j=0}^{2q-2}\frac{\bar{R}_j\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\left(\frac{cA}{L}\bar{\xi}\right)^j}+\frac{\bar{R}^{\text{rem}}_{2q-1}\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\left(\frac{cA}{L}\bar{\xi}\right)^{2q-1}}\right]}{\bar{\xi}^{\ell+\frac{1}{2}}}\right\} \\ &\times e^{i\left\{kL\left[1+\sin(\theta)\right]+2\pi(n+m)\right\}\frac{cA}{L}\bar{\xi}}\,d\bar{\xi}.\end{aligned} \tag{3.179}$$

This reformulation contains an additional factor of $\left(\frac{A}{L}\right)^{-1}$ relative to (3.169) due to the integration of the periodic factor $e^{i\left\{kL\left[1+\sin(\theta)\right]+2\pi(n+m)\right\}\frac{cA}{L}\bar{\xi}}$.

Now, the integral in (3.179) is absolutely convergent: the term with $\frac{1}{\bar{\xi}^{\ell+\frac{3}{2}}}$ decays sufficiently rapidly for $\ell=0,1,2,\ldots,p-1$, and the term with $\frac{1}{\bar{\xi}^{\ell+\frac{1}{2}}}$ has finite support due to the derivative of $\bar{P}_2\left[c,\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)\right]$. Since these terms are $\mathcal{O}(1)$ in $\frac{A}{L}$ as $A\to\infty$, we

therefore have

$$
\begin{aligned}
|I_{n,\ell,m}(x,k,c,A)| &\leq \frac{|d_n|\left|v_{n,\ell,m}\left(kL,\frac{2\pi x}{L}\right)\right|}{|kL\left[1+\sin(\theta)\right]+2\pi(n+m)|}\left(\frac{cA}{L}\right)^{-\ell-\frac{1}{2}} \\
&\quad \times \int_{\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty} \left| \frac{\left(-\ell-\frac{1}{2}\right)\bar{P}_2\left[c,\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)\right]}{\bar{\xi}^{\ell+\frac{3}{2}}} \right. \\
&\quad \left. + \frac{-\left.\frac{\partial S\left(z,1,\frac{1}{c}\right)}{\partial z}\right|_{z=\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)}\left[\sum_{j=0}^{2q-2}\frac{\bar{R}_j\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\left(\frac{cA}{L}\bar{\xi}\right)^j}+\frac{\bar{R}_{2q-1}^{\mathrm{rem}}\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\left(\frac{cA}{L}\bar{\xi}\right)^{2q-1}}\right]}{\bar{\xi}^{\ell+\frac{1}{2}}} \right| d\bar{\xi} \\
&\leq \frac{|d_n|\left|v_{n,\ell,m}\left(kL,\frac{2\pi x}{L}\right)\right| M_{\ell,q}(x,c)}{|kL\left[1+\sin(\theta)\right]+2\pi(n+m)|}\left(\frac{A}{L}\right)^{-\ell-\frac{1}{2}},
\end{aligned}
\tag{3.180}
$$

where $M_{\ell,q}(x,c)$ is some positive real function of x,c that is independent of A,n,m.

For $p=1$, we employ the above analysis for $I_{n,0,m}(x,k,c,A)$, $m\in\mathbb{Z}$ to determine that they all (along with $I_n^{\mathrm{rem}}(x,k,c,A)$) are of size $\mathcal{O}\left(\left(\frac{A}{L}\right)^{-p+\frac{1}{2}}\right)$ as $A\to\infty$. For $p>1$, however, we need to repeat the process of making Fourier expansions and integrating by parts if $p-\ell>1$.

Our general procedure if $p-\ell>1$ is as follows: we write the sum multiplying the factor $\frac{1}{\bar{\xi}^{\ell+\frac{1}{2}}}$ in (3.179) as

$$
\begin{aligned}
&\frac{1}{\left(\frac{cA}{L}\bar{\xi}\right)^{\ell+\frac{1}{2}}}\left[\sum_{j=0}^{2q-2}\frac{\tilde{R}_j\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\left(\frac{cA}{L}\bar{\xi}\right)^j}+\frac{\tilde{R}_{2q-1}^{\mathrm{rem}}\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\left(\frac{cA}{L}\bar{\xi}\right)^{2q-1}}\right] \\
&\qquad = \frac{1}{\left(\frac{cA}{L}\bar{\xi}\right)^{p-\frac{1}{2}}}\left[\sum_{j=p-\ell-1}^{2q-2}\frac{\tilde{R}_j\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\left(\frac{cA}{L}\bar{\xi}\right)^{j-p+\ell+1}}+\frac{\tilde{R}_{2q-1}^{\mathrm{rem}}\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\left(\frac{cA}{L}\bar{\xi}\right)^{2q-p+\ell}}\right] \\
&\qquad + \frac{1}{\left(\frac{cA}{L}\bar{\xi}\right)^{\ell+\frac{1}{2}}}\sum_{j=0}^{p-\ell-2}\frac{\tilde{R}_j\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\left(\frac{cA}{L}\bar{\xi}\right)^j},
\end{aligned}
\tag{3.181}
$$

recalling that q was chosen sufficiently large relative to p (Remark 3.1.5). For the first sum on the right-hand side, it immediately holds that the integrals containing these terms are all absolutely convergent and are of size $\mathcal{O}\left(\left(\frac{A}{L}\right)^{-p+\frac{1}{2}}\right)$ as $A\to\infty$. For the second sum on the right-hand side, however, we expand the periodic factors into uniformly con-

vergent Fourier series in $\bar{\xi}$ as before, and the resulting terms (in addition to the term $\frac{1}{\left(\frac{cA}{L}\right)^{\ell+\frac{1}{2}}}\frac{\left(-\ell-\frac{1}{2}\right)\bar{P}_2\left[c,\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)\right]}{\bar{\xi}^{\ell+\frac{3}{2}}}$) are distributed so that we have integrals multiplied by the factors $\frac{1}{\left(\frac{cA}{L}\right)^{j-\frac{1}{2}}}$, $j=\ell+1,\ell+2,\ell+3,\ldots,p-1$. Integration by parts is then employed: as was done for (3.169), the smoothly decaying terms are differentiated and the periodic factor $e^{i\{kL[1+\sin(\theta)]+2\pi(n+m+r)\}\frac{cA}{L}\bar{\xi}}$ ($r\in\mathbb{Z}$ comes from the new Fourier expansions) is integrated, which generates an additional $\left(\frac{A}{L}\right)^{-1}$ factor in our estimates of the integrals. This process of Fourier expansions and integrations by parts is continued until all of the resulting integrals are of size $\mathcal{O}\left(\left(\frac{A}{L}\right)^{-p+\frac{1}{2}}\right)$ as $A\to\infty$. Finally, we absolutely sum the series of integrals that are generated by this procedure.

As an example of this procedure, we consider the term $\frac{1}{\left(\frac{cA}{L}\bar{\xi}\right)^{\ell+\frac{1}{2}}}\tilde{R}_0\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)$ in (3.181). The part of the function $I_{n,\ell,m}(x,k,c,A)$ (3.179) containing this term is

$$\begin{aligned}\tilde{I}_{n,\ell,m}(x,k,c,A)\equiv&\frac{d_n e^{in\frac{2\pi x}{L}}v_{n,\ell,m}\left(kL,\frac{2\pi x}{L}\right)}{i\left\{kL\left[1+\sin(\theta)\right]+2\pi(n+m)\right\}}\frac{1}{\left(\frac{cA}{L}\right)^{\ell+\frac{1}{2}}}\\&\times\int_{\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty}\frac{\left.\frac{\partial S\left(z,1,\frac{1}{c}\right)}{\partial z}\right|_{z=\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)}\tilde{R}_0\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)}{\bar{\xi}^{\ell+\frac{1}{2}}}\\&\times e^{i\{kL[1+\sin(\theta)]+2\pi(n+m)\}\frac{cA}{L}\bar{\xi}}\,d\bar{\xi}.\end{aligned}\tag{3.182}$$

The periodic function $\tilde{R}_0\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)$ has the Fourier expansion

$$\tilde{R}_0\left(\frac{2\pi x}{L},2\pi\frac{cA}{L}\bar{\xi}\right)=\sum_{r=-\infty}^{\infty}\tilde{v}_r\left(\frac{2\pi x}{L}\right)e^{i2\pi r\frac{cA}{L}\bar{\xi}},\tag{3.183}$$

which is a series that converges uniformly for $(x,\bar{\xi})\in[0,L]\times\left[\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right),\infty\right)$ with coefficients $\tilde{v}_r$ that (in magnitude) decrease super-algebraically in r as $r\to\pm\infty$ due to the smoothness of the grating profile. Therefore, we may write

$$\tilde{I}_{n,\ell,m}(x,k,c,A)=\sum_{r=-\infty}^{\infty}\tilde{I}_{n,\ell,m,r}(x,k,c,A),\tag{3.184}$$

where

$$\tilde{I}_{n,\ell,m,r}(x,k,c,A) \equiv \frac{d_n e^{in\frac{2\pi x}{L}} v_{n,\ell,m}\left(kL,\frac{2\pi x}{L}\right)\tilde{v}_r\left(\frac{2\pi x}{L}\right)}{i\left\{kL\left[1+\sin(\theta)\right]+2\pi(n+m)\right\}} \frac{1}{\left(\frac{cA}{L}\right)^{\ell+\frac{1}{2}}} \times \int_{\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty} \frac{\left.\frac{\partial S\left(z,1,\frac{1}{c}\right)}{\partial z}\right|_{z=\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{cA}{L}\bar{\xi}\right)}}{\bar{\xi}^{\ell+\frac{1}{2}}} e^{i\left\{kL\left[1+\sin(\theta)\right]+2\pi(n+m+r)\right\}\frac{cA}{L}\bar{\xi}}\, d\bar{\xi}. \tag{3.185}$$

Thus, integrating by parts and bounding $\left|\tilde{I}_{n,\ell,m,r}(x,k,c,A)\right|$ by the integral of the absolute value of the integrand (as we did before) results in the relation

$$\begin{aligned}&\left|\tilde{I}_{n,\ell,m,r}(x,k,c,A)\right| \\ &\leq \frac{|d_n|\left|v_{n,\ell,m}\left(kL,\frac{2\pi x}{L}\right)\right|\left|\tilde{v}_r\left(\frac{2\pi x}{L}\right)\right|\tilde{M}_\ell(x,c)}{\left|kL\left[1+\sin(\theta)\right]+2\pi(n+m)\right|\left|kL\left[1+\sin(\theta)\right]+2\pi(n+m+r)\right|}\left(\frac{A}{L}\right)^{-\ell-\frac{3}{2}},\end{aligned} \tag{3.186}$$

where $\tilde{M}_\ell(x,c)$ is some positive real function of x,c that is independent of A,n,m,r. Finally, since the coefficients $\tilde{v}_r$ (in magnitude) decrease super-algebraically in r as $r\to\pm\infty$, it follows that

$$\begin{aligned}&\max_{x\in[0,L]}\left|\tilde{I}_{n,\ell,m}(x,k,c,A)\right| \\ &\quad=\max_{x\in[0,L]}\left|\sum_{r=-\infty}^{\infty}\tilde{I}_{n,\ell,m,r}(x,k,c,A)\right| \\ &\quad\leq\sum_{r=-\infty}^{\infty}\max_{x\in[0,L]}\left|\tilde{I}_{n,\ell,m,r}(x,k,c,A)\right| \\ &\quad\leq\max_{x\in[0,L]}\frac{|d_n|\left|v_{n,\ell,m}\left(kL,\frac{2\pi x}{L}\right)\right|\tilde{M}_\ell(x,c)}{\left|kL\left[1+\sin(\theta)\right]+2\pi(n+m)\right|}\left(\frac{A}{L}\right)^{-\ell-\frac{3}{2}}\sum_{r=-\infty}^{\infty}\frac{\left|\tilde{v}_r\left(\frac{2\pi x}{L}\right)\right|}{\left|kL\left[1+\sin(\theta)\right]+2\pi(n+m+r)\right|} \\ &\quad=\max_{x\in[0,L]}\frac{|d_n|\left|v_{n,\ell,m}\left(kL,\frac{2\pi x}{L}\right)\right|\tilde{M}_\ell(x,c)\tilde{N}_{n,m}(x,k)}{\left|kL\left[1+\sin(\theta)\right]+2\pi(n+m)\right|}\left(\frac{A}{L}\right)^{-\ell-\frac{3}{2}},\end{aligned} \tag{3.187}$$

where

$$\tilde{N}_{n,m}(x,k)\equiv\sum_{r=-\infty}^{\infty}\frac{\left|\tilde{v}_r\left(\frac{2\pi x}{L}\right)\right|}{\left|kL\left[1+\sin(\theta)\right]+2\pi(n+m+r)\right|} \tag{3.188}$$

is a positive real function of x,k that is independent of A,r; we note that the functions

$\tilde{N}_{n,m}(x,k)$ have the bound

$$\tilde{N}_{n,m}(x,k) \leq \bar{N}(x,k) \tag{3.189}$$

for all $n,m \in \mathbb{Z}$ for some positive real function $\bar{N}(x,k)$ that is independent of n,m, since (by assumption) $|kL\,[1+\sin(\theta)]+2\pi n| > \eta > 0$ for every $n \in \mathbb{Z}$. Thus, if $p-\ell=2$, then we have shown that

$$\max_{x\in[0,L]} \left|\tilde{I}_{n,\ell,m}(x,k,c,A)\right| = \mathcal{O}\left(\left(\frac{A}{L}\right)^{-p+\frac{1}{2}}\right) \tag{3.190}$$

as $A \to \infty$. Otherwise, we again make Fourier expansions and use integration by parts in order to generate yet another $\left(\frac{A}{L}\right)^{-1}$ factor in our estimate, and we do this a finite number of times until all of the resulting integrals (which are then absolutely summed) are of size $\mathcal{O}\left(\left(\frac{A}{L}\right)^{-p+\frac{1}{2}}\right)$ as $A \to \infty$.

This procedure—making Fourier expansions and performing integrations by parts until all of the resulting integrals are of size $\mathcal{O}\left(\left(\frac{A}{L}\right)^{-p+\frac{1}{2}}\right)$ as $A \to \infty$ and then absolutely summing the series of integrals—results in the estimate

$$|I_{n,\ell,m}(x,k,c,A)| \leq \frac{|d_n|\left|v_{n,\ell,m}\left(kL,\frac{2\pi x}{L}\right)\right| N_{n,\ell,m,q}(x,k,c)}{|kL\,[1+\sin(\theta)]+2\pi(n+m)|}\left(\frac{A}{L}\right)^{-p+\frac{1}{2}}, \tag{3.191}$$

where $N_{n,\ell,m,q}(x,k,c)$ is some positive real function of x,k,c that is independent of A; as with $M_{\ell,q}(x,c)$ of (3.180) and $\tilde{M}_\ell(x,c)\tilde{N}_{n,m}(x,k)$ of (3.187), the functions $N_{n,\ell,m,q}(x,k,c)$ are bounded from above by functions that are independent of n,m, i.e.,

$$N_{n,\ell,m,q}(x,k,c) \leq \bar{N}_{\ell,q}(x,k,c) \tag{3.192}$$

for all $n,m \in \mathbb{Z}$ for certain positive real functions $\bar{N}_{\ell,q}(x,k,c)$ that are independent of n,m (we recall that q is a constant that is chosen sufficiently large relative to p). We note that estimates of this form hold for every integer $p \geq 1$, where p is the order of the asymptotic expansion (3.156).

Summation of individual estimates. Now, the functions $\frac{\left|v_{n,\ell,m}\left(kL,\frac{2\pi x}{L}\right)\right|N_{n,\ell,m,q}(x,k,c)}{|kL[1+\sin(\theta)]+2\pi(n+m)|}$ decrease super-algebraically in m as $m \to \pm\infty$, as can be seen by the bounds (3.167) and (3.192) for the Fourier coefficients $v_{n,\ell,m}$ and the functions $N_{n,\ell,m,q}(x,k,c)$, respectively.

Thus, we have

$$\begin{aligned}
&\max_{x\in[0,L]} |I_n(x,k,c,A)| \\
&\quad = \max_{x\in[0,L]} \left| \sum_{\ell=0}^{p-1} \sum_{m=-\infty}^{\infty} I_{n,\ell,m}(x,k,c,A) + I_n^{\text{rem}}(x,k,c,A) \right| \\
&\quad \leq \sum_{\ell=0}^{p-1} \sum_{m=-\infty}^{\infty} \max_{x\in[0,L]} |I_{n,\ell,m}(x,k,c,A)| + \max_{x\in[0,L]} |I_n^{\text{rem}}(x,k,c,A)| \\
&\quad \leq |d_n| \left\{ \sum_{\ell=0}^{p-1} \sum_{m=-\infty}^{\infty} \max_{x\in[0,L]} \frac{\left| v_{n,\ell,m}\left(kL, \frac{2\pi x}{L}\right) \right| N_{n,\ell,m,q}(x,k,c)}{|kL\left[1+\sin(\theta)\right] + 2\pi(n+m)|} + \max_{x\in[0,L]} M_{n,p}^{\text{rem}}(x,k,c) \right\} \left(\frac{A}{L}\right)^{-p+\frac{1}{2}}
\end{aligned} \tag{3.193}$$

as $A \to \infty$ for every integer $p \geq 1$. As stated previously, the factors in this inequality which multiply $\left(\frac{A}{L}\right)^{-p+\frac{1}{2}}$ are all independent of A.

Finally, we see from the bounds (3.96), (3.167), (3.175) and (3.192) that the estimates (3.193) for $\max_{x\in[0,L]} |I_n(x,k,c,A)|$ decrease super-algebraically in n as $n \to \pm\infty$. Therefore, they can be summed over all n, and we conclude that

$$\begin{aligned}
\max_{x\in[0,L]} \left|I_+^2(x,k,c,A)\right| &= \max_{x\in[0,L]} \left| \sum_{n=-\infty}^{\infty} I_n(x,k,c,A) \right| \\
&\leq \sum_{n=-\infty}^{\infty} \max_{x\in[0,L]} |I_n(x,k,c,A)| \\
&= \mathcal{O}\left(\left(\frac{A}{L}\right)^{-p+\frac{1}{2}} \right)
\end{aligned} \tag{3.194}$$

as $A \to \infty$ for every integer $p \geq 1$.

Remark 3.1.6. *For $I_+^1(x,k,c,A)$, the only necessary change to the proof above is that we modify (3.89) to be*

$$\bar{\phi}_+(x,t) \equiv t\sqrt{1+\left(\frac{h}{2L}\right)^2 \left\{ \frac{\bar{f}\left[2\pi\left(\frac{x}{L}+t\right)\right] - \bar{f}\left(\frac{2\pi x}{L}\right)}{t} \right\}^2} + t\sin(\theta) \tag{3.195}$$

based upon the formula (2.91) for $\phi_1(x,x')$. It still holds that $\frac{\partial \bar{\phi}_+}{\partial t}$ has an upper and lower bound for $t \in (0,\infty)$. Thus, we still may choose some constant $c_1 \geq 0$ such that

$$\frac{\partial}{\partial t}\left[\bar{\phi}_+(x,t) + c_1 t\right] > \epsilon > 0 \tag{3.196}$$

for all $t > 0$, so we again may make the change of variables

$$\xi \equiv \tilde{\phi}_+(x,t) \equiv \frac{1}{1+\sin(\theta)+c_1}\left[\bar{\phi}_+(x,t)+c_1 t\right] \tag{3.197}$$

with the corresponding inverse

$$\tilde{\phi}_+^{-1}(x,\xi) \equiv t. \tag{3.198}$$

The various expansions that follow once again lead to integrals $I_n(x,k,c,A)$ of the form (3.162), and integration by parts works in the same way as before due to the assumption $|kL[1+\sin(\theta)]+2\pi n| > \eta > 0$ for every $n \in \mathbb{Z}$. Thus, the estimation of the size in $\frac{A}{L}$ of $I_+^1(x,k,c,A)$ also holds.

Part 2: To prove this part, we verify that the "preliminary results" for $I_-^2(x,k,c,A)$ are similar to those established in Section 3.1.3.1 for $I_+^2(x,k,c,A)$ and indicate how the proof therefore follows closely to the one given in Part 1.

We use $P_2(x,x+x',c,A)=0$ for $x' \in [-cA,0]$ and $t \equiv -\frac{x'}{L}$ to write

$$I_-^2(x,k,c,A) = \int_{\frac{cA}{L}}^{\infty} \bar{P}_2\left(c,\frac{Lt}{cA}\right)\bar{g}_-(x,t)\,\frac{h\left[kL\bar{u}_-(x,t)\right]}{e^{ikL\bar{u}_-(x,t)}}e^{ikL\bar{\phi}_-(x,t)}\bar{\varphi}\left[2\pi\left(\frac{x}{L}-t\right)\right]\,dt. \tag{3.199}$$

Also, given (3.83), we let

$$\bar{g}_-(x,t) \equiv \frac{i}{2}\frac{h}{2L}\frac{1}{t}\frac{\frac{\bar{f}\left[2\pi\left(\frac{x}{L}-t\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)}{t}+2\pi\bar{f}'\left(\frac{2\pi x}{L}\right)}{1+\left(\frac{h}{2L}\right)^2\left\{\frac{\bar{f}\left[2\pi\left(\frac{x}{L}-t\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)}{t}\right\}^2}, \tag{3.200}$$

$$\bar{u}_-(x,t) \equiv t\sqrt{1+\left(\frac{h}{2L}\right)^2\left\{\frac{\bar{f}\left[2\pi\left(\frac{x}{L}-t\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)}{t}\right\}^2} \tag{3.201}$$

and

$$\begin{aligned}\bar{\phi}_-(x,t) \equiv & t\sqrt{1+\left(\frac{h}{2L}\right)^2\left\{\frac{\bar{f}\left[2\pi\left(\frac{x}{L}-t\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)}{t}\right\}^2} \\ & -(\sin(\theta),\cos(\theta))\cdot\left(t,\frac{h}{2L}\left\{\bar{f}\left[2\pi\left(\frac{x}{L}-t\right)\right]-\bar{f}\left(\frac{2\pi x}{L}\right)\right\}\right).\end{aligned} \tag{3.202}$$

Noting that

$$\frac{\partial \bar{\phi}_{-}(x,t)}{\partial t} = -L \left. \frac{\partial \phi_2(x, x+x')}{\partial x'} \right|_{x'=-Lt}, \tag{3.203}$$

we can—as we did with $\frac{\partial \bar{\phi}_{+}}{\partial t}$—choose a constant $c_1 \geq 0$ such that

$$\frac{\partial}{\partial t}\left[\bar{\phi}_{-}(x,t) + c_1 t\right] > \epsilon > 0. \tag{3.204}$$

Thus, we make the change of variables

$$\xi \equiv \tilde{\phi}_{-}(x,t), \tag{3.205}$$

where $\tilde{\phi}_{-}(x,t)$ is given by

$$\tilde{\phi}_{-}(x,t) \equiv \frac{1}{1-\sin(\theta)+c_1}\left[\bar{\phi}_{-}(x,t) + c_1 t\right], \tag{3.206}$$

and we define the inverse

$$\tilde{\phi}_{-}^{-1}(x,\xi) \equiv t. \tag{3.207}$$

Remark 3.1.7. *Here, as with* $\frac{\partial \bar{\phi}_{+}}{\partial t}$ *(Remark 3.1.4), we may let* $c_1 = 0$ *for simple-reflection cases, since for such problems we have* $\frac{\partial \phi(x,x')}{\partial x'} < 0$ *for all* $x' < x$ *(Corollary 2.3.1). But, we must choose* $c_1 > 0$ *for multiple-reflection cases.*

Therefore, writing

$$I_{-}^{2}(x,k,c,A) = \int_{\tilde{\phi}_{-}\left(x,\frac{cA}{L}\right)}^{\infty} \bar{P}_2\left(c, \frac{L\tilde{\phi}_{-}^{-1}}{cA}\right) \bar{g}_{-}\left(x, \tilde{\phi}_{-}^{-1}\right) \frac{h\left(kL\bar{u}_{-}\left(x,\tilde{\phi}_{-}^{-1}\right)\right)}{e^{ikL\bar{u}_{-}\left(x,L,h,\tilde{\phi}_{-}^{-1}\right)}} e^{ikL[1-\sin(\theta)+c_1]\xi} \times e^{-ikLc_1\tilde{\phi}_{-}^{-1}} \bar{\varphi}\left[2\pi\left(\frac{x}{L} - \tilde{\phi}_{-}^{-1}\right)\right] \frac{\partial \tilde{\phi}_{-}^{-1}}{\partial \xi}\, d\xi, \tag{3.208}$$

the proof for $I_{-}^{2}(x,k,c,A)$ can be completed in a manner analogous to the proof for $I_{+}^{2}(x,k,c,A)$ in Part 1. As indicated by the phase terms in the above integrand, the assumption $|kL[1-\sin(\theta)] + 2\pi n| > \eta > 0$ for every $n \in \mathbb{Z}$ must hold in order to employ integration by parts in the way described in Part 1.

Just as the proof for $I_{+}^{1}(x,k,c,A)$ follows closely to that for $I_{+}^{2}(x,k,c,A)$, with the only difference being the use of $\phi_1(x,x')$ instead of $\phi_2(x,x')$ (see Remark 3.1.6), the proof for

$I^1_-(x,k,c,A)$ is similar to that for $I^2_-(x,k,c,A)$.

Part 3: The result follows from Parts 1 and 2 by the inequality

$$\max_{x\in[0,L]} |I^m(x,k,c,A)| \leq \max_{x\in[0,L]} \left|I^m_+(x,k,c,A)\right| + \max_{x\in[0,L]} \left|I^m_-(x,k,c,A)\right|. \tag{3.209}$$

□

3.1.3.4 Super-Algebraic Convergence of Solutions for Smooth Gratings

We conclude our discussion about the properties of the solutions $\mu^A_m(x)$ by stating the following key characteristic of our algorithm.

Theorem 3.1.10. *Let the grating profile $f(x)$ belong to $C^\infty_{per}(L)$. Then, for k which are not Wood Anomaly values and sufficiently large A,*

1. *the solutions $\mu^A_m(x)$ of the approximating integral equations (3.1) exist;*
2. *the solutions $\mu^A_m(x)$ as well as the solutions $\mu_m(x)$ of the exact integral equations (3.5) belong to $C^\infty_{per}(L)$;*
3. *the functions $\mu^A_m(x)$ converge super-algebraically in A, uniformly for $x \in [0,L]$, to $\mu_m(x)$ as $A \to \infty$.*

Proof. Part 1 is a result of Theorem 3.1.4. Part 2 is the result of Theorems 3.1.7 and 3.1.8. Part 3 follows from the error estimate (3.42) of Theorem 3.1.4 and the super-algebraic convergence result of Theorem 3.1.9. □

3.2 Properties of Numerical Method

Using the parameterization $x(t) \equiv \frac{L}{2\pi}t$ and setting $\bar{\mu}^A_m(t) \equiv \mu^A_m[x(t)]$ ($\bar{\mu}^A_m \in C^0_{\text{per}}(2\pi)$), we re-write the approximating integral equations (2.132) as

$$\bar{\mu}^A_m(t) \pm \frac{L}{2\pi}\int_{t-\frac{2\pi A}{L}}^{t+\frac{2\pi A}{L}} K^m(t,\tau)\bar{\mu}^A_m(\tau)\,d\tau = q_m[x(t)], \quad 0 \leq t \leq 2\pi \tag{3.210}$$

($m = 1$ or $m = 2$), where $K^m(t,\tau)$ is given by (2.142). As described in Section 2.5.2, splitting $K^m(t,\tau)$ by using certain smooth windowing functions and series expansions in

order to isolate its logarithmic singularity at $\tau = t$ and then applying certain spectrally accurate quadrature formulas leads to the numerical approximation equations

$$\begin{aligned}\bar{\mu}_m^{A,n_\mathrm{i}}(t) \pm \frac{L}{2\pi} \sum_{\left\{j:\, |t-t_j|<\frac{2\pi A_\mathrm{sp}}{L}\right\}} \left\{ R_j^{\left(\frac{n_\mathrm{i}}{2}\right)}(t) K_{\mathrm{sp1}}^m(t,t_j) + \frac{2\pi}{n_\mathrm{i}} K_{\mathrm{sp2}}^m(t,t_j) \right\} \bar{\mu}_m^{A,n_\mathrm{i}}(t_j) \\ \pm \frac{L}{2\pi}\frac{2\pi}{n_\mathrm{i}} \sum_{\left\{j:\, \frac{2\pi c_\mathrm{sp} A_\mathrm{sp}}{L}<|t-t_j|<\frac{2\pi A}{L}\right\}} K_{\mathrm{rp}}^m(t,t_j) \bar{\mu}_m^{A,n_\mathrm{i}}(t_j) = q_m\left[x(t)\right], \quad 0 \le t \le 2\pi,\end{aligned} \tag{3.211}$$

where t_j are the integration points (2.159) and $\bar{\mu}_m^{A,n_\mathrm{i}}(t)$ are the approximations of $\bar{\mu}_m^A(t)$ that arise from using the quadrature rule (2.164) with n_i integration points per period. Discretizing these equations results in the linear systems

$$\tilde{\mu}_\ell \pm \sum_{j=0}^{n_\mathrm{i}-1} \hat{a}_{\ell,j}\hat{\mu}_j = q_m\left[x(t_{\ell\cdot n_\mathrm{mult}})\right], \quad \ell = 0,1,2,\ldots,n_\mathrm{t}-1 \tag{3.212}$$

(Section 2.5.3). Here $\tilde{\mu}_\ell,\ \ell = 0,1,2,\ldots,n_\mathrm{t}-1$ are the approximate values of $\bar{\mu}_m^{A,n_\mathrm{i}}(t)$ at the n_t "target points" $t = t_{\ell\cdot n_\mathrm{mult}}$ lying in $[0,2\pi)$, $\hat{\mu}_j$ are the values of the Fourier interpolation of $\tilde{\mu}_\ell,\ \ell = 0,1,2,\ldots,n_\mathrm{t}-1$ onto $t_j,\ j = 0,1,2,\ldots,n_\mathrm{i}-1$ (these values of $\hat{\mu}_j$ are periodically extended to all of the integration points), $\sum_{j=0}^{n_\mathrm{i}-1}\hat{a}_{\ell,j}\hat{\mu}_j$ is given by (2.167) and $n_\mathrm{i} = n_\mathrm{t}\times n_\mathrm{mult}$ for some positive integer n_mult; we note that

$$\tilde{\mu}_\ell = \hat{\mu}_\ell = \bar{\mu}_m^{A,n_\mathrm{i}}\left(t_\ell\right), \quad \ell = 0,1,2,\ldots,n_\mathrm{t}-1 \tag{3.213}$$

if $n_\mathrm{i} = n_\mathrm{t}$.

In Section 3.2.1, we show that the quadratures converge in n_i to the integral operators as $n_\mathrm{i}\to\infty$ and thus that the solutions $\bar{\mu}_m^{A,n_\mathrm{i}}(t)$ also converge in n_i, uniformly for $t\in[0,2\pi]$, to $\bar{\mu}_m^A(t)$ as $n_\mathrm{i}\to\infty$. In particular, this convergence is super-algebraic for cases with smooth scattering surfaces. Therefore, if we set $n_\mathrm{i} = n_\mathrm{t}$—i.e., set $n_\mathrm{mult} = 1$ so that the linear systems (3.212) are the ones used in the classical Nyström approach—the values $\tilde{\mu}_\ell,\ \ell = 0,1,2,\ldots,n_\mathrm{t}-1$ converge uniformly to $\bar{\mu}_m^A\left(t_\ell\right),\ \ell = 0,1,2,\ldots,n_\mathrm{t}-1$ as $n_\mathrm{t}\to\infty$ (super-algebraically when the grating profile is C^∞).

Remark 3.2.1. *As discussed earlier (Section 2.5.3), for certain problems we choose n_{mult} to be greater than 1 for the sake of computational efficiency. We do not develop error*

bounds for our approximations when $n_{mult} > 1$. We do, however, note that the values $\tilde{\mu}_\ell,\ \ell = 0, 1, 2, \ldots, n_t - 1$ do not converge to $\bar{\mu}_m^A(t_\ell),\ \ell = 0, 1, 2, \ldots, n_t - 1$ as $n_i \to \infty$ if the number n_t of target points is kept fixed, yet increasing values of n_{mult} (and thus n_i) can produce increasingly accurate computations up to a desired level of accuracy (e.g., machine precision) for sufficiently large fixed n_t. See Section 4.1 for numerical results illustrating these points.

After the convergence proofs of Section 3.2.1, we demonstrate in Section 3.2.2 that for problems in which only simple reflections arise our numerical method requires $\mathcal{O}(1)$ total computational time as $k \to \infty$ (for a given grating profile $f(x)$ and incidence angle θ) in order to compute the values $\tilde{\mu}_\ell,\ \ell = 0, 1, 2, \ldots, n_t - 1$ to arbitrary accuracy.

3.2.1 Convergence in Mesh Size

As stated previously, we reformulate the kernel $K^m(t,\tau)$ (2.142) using the smooth windowing functions $P_{\text{sp}}[x(t), x(\tau), c_{\text{sp}}, A_{\text{sp}}]$ (2.145) and $P_{\text{rp}}[x(t), x(\tau), c_{\text{sp}}, A_{\text{sp}}, c, A]$ (2.146):

$$K^m(t,\tau) = K^m_{\text{sp}}(t,\tau) + K^m_{\text{rp}}(t,\tau), \tag{3.214}$$

where $K^m_{\text{sp}}(t,\tau)$ and $K^m_{\text{rp}}(t,\tau)$ are given by (2.147) and (2.148), respectively. We recall that the support of $K^m_{\text{sp}}(t,\tau)$ lies within a subinterval of $(t-2\pi, t+2\pi)$, because $P_{\text{sp}}[x(t), x(\tau), c_{\text{sp}}, A_{\text{sp}}]$ is centered about $\tau = t$ and $A_{\text{sp}} < L$. Thus, an application of certain series expansions for the first and second kind Bessel functions $J_1(x)$ and $Y_1(x)$ (see Remark 2.5.2) results in the splitting

$$K^m(t,\tau) = K^m_{\text{sp1}}(t,\tau)\log\left[4\sin^2\left(\frac{t-\tau}{2}\right)\right] + K^m_{\text{sp2}}(t,\tau) + K^m_{\text{rp}}(t,\tau); \tag{3.215}$$

here $K^m_{\text{sp1}}(t,\tau)$ and $K^m_{\text{sp2}}(t,\tau)$ are given by (2.150) and (2.151), respectively, and for smooth grating profiles $f[x(t)]$ this formula has the property that $K^m_{\text{rp}}(t,\tau)$, $K^m_{\text{sp1}}(t,\tau)$ and $K^m_{\text{sp2}}(t,\tau)$ are all C^∞ (with finite support) in τ and $C^\infty_{\text{per}}(2\pi)$ in t.

For the numerical approximation equations (3.211), we first demonstrate that the quadratures converge in n_{i} as $n_{\text{i}} \to \infty$ (Section 3.2.1.1), and then we show how this leads to the convergence in n_{i} of the solutions $\bar{\mu}_m^{A,n_{\text{i}}}(t)$ as $n_{\text{i}} \to \infty$ (Section 3.2.1.2). In Section 3.2.1.3, we prove that the convergence of the quadratures (and thus the convergence of the solutions)

is super-algebraic for cases with C^∞ gratings.

3.2.1.1 Convergence of Quadratures

Following the terminology of [35], we say that a set of linear operators on $C^0_{\text{per}}(2\pi)$ is "collectively compact" if, for each bounded subset of $C^0_{\text{per}}(2\pi)$, the image set is relatively compact. Also, we say that a sequence of linear operators $(\mathcal{A}_n)$ on $C^0_{\text{per}}(2\pi)$ is "pointwise convergent" to the linear operator $\mathcal{A}$ on $C^0_{\text{per}}(2\pi)$ if

$$||(\mathcal{A} - \mathcal{A}_n)\,\varphi||_\infty \to 0 \tag{3.216}$$

as $n \to \infty$ for every $\varphi \in C^0_{\text{per}}(2\pi)$.

Given these definitions and the quadrature rule (2.164), we prove the following theorem.

Theorem 3.2.1. *The sequences of numerical integration operators*

$$\begin{aligned}\mathcal{A}^m_{n_i}\varphi(t) \equiv \frac{L}{2\pi} \sum_{\left\{j:\,|t-t_j|<\frac{2\pi A_{sp}}{L}\right\}} \left\{ R_j^{\left(\frac{n_i}{2}\right)}(t) K^m_{sp1}(t,t_j) + \frac{2\pi}{n_i} K^m_{sp2}(t,t_j) \right\} \varphi(t_j) \\ + \frac{L}{2\pi}\frac{2\pi}{n_i} \sum_{\left\{j:\,\frac{2\pi c_{sp} A_{sp}}{L}<|t-t_j|<\frac{2\pi A}{L}\right\}} K^m_{rp}(t,t_j)\varphi(t_j)\end{aligned} \tag{3.217}$$

which approximate

$$\mathcal{A}^m\varphi(t) \equiv \frac{L}{2\pi}\int_{t-\frac{2\pi A}{L}}^{t+\frac{2\pi A}{L}} K^m(t,\tau)\varphi(\tau)\,d\tau \tag{3.218}$$

for $\varphi \in C^0_{per}(2\pi)$ are collectively compact and pointwise convergent.

This result is established by individually considering each of the quadrature formulas (2.160)–(2.163).

Theorem 3.2.2. *The sequences of numerical integration operators*

$$\mathcal{A}^m_{n_i}\varphi(t) \equiv \frac{L}{2\pi} \sum_{\left\{j:\,|t-t_j|<\frac{2\pi A_{sp}}{L}\right\}} R_j^{\left(\frac{n_i}{2}\right)}(t) K^m_{sp1}(t,t_j)\varphi(t_j) \tag{3.219}$$

on $C^0_{per}(2\pi)$ which approximate

$$\mathcal{A}^m\varphi(t) \equiv \frac{L}{2\pi}\int_{t-\frac{2\pi A_{sp}}{L}}^{t+\frac{2\pi A_{sp}}{L}} K^m_{sp1}(t,\tau)\log\left[4\sin^2\left(\frac{t-\tau}{2}\right)\right]\varphi(\tau)\,d\tau \tag{3.220}$$

for $\varphi \in C^0_{per}(2\pi)$ are collectively compact and pointwise convergent.

Proof. Since $P_{\text{sp}}\left[x(t), x(\tau), c_{\text{sp}}, A_{\text{sp}}\right]$ decays smoothly to 0 within a subinterval of $(t-2\pi, t+2\pi)$, it follows that $K^m_{\text{sp1}}(t,\tau)$ can be periodically extended in τ (with period 4π) for each $t \in [0, 2\pi]$. Also, $K^m_{\text{sp1}}(t,\tau)$ are 2π-periodic in t, and $K^m_{\text{sp1}}(t,\tau) = 0$ in neighborhoods of $\tau = t \pm 2\pi$. Thus, we may apply the quadrature formula (2.137)—with an appropriately modified interval of integration—to $\mathcal{A}^m\varphi(t)$, which results in $\mathcal{A}^m_{n_{\text{i}}}\varphi(t)$. This quadrature formula is convergent, and the weights $R_j^{\left(\frac{n_{\text{i}}}{2}\right)}(t)$ (2.138) satisfy

$$\lim_{\tau \to t} \sup_{\frac{n_{\text{i}}}{2} \in \mathbb{N}} \sum_{\left\{j:\, |t-t_j| < \frac{2\pi A_{\text{sp}}}{L}\right\}} \left| R_j^{\left(\frac{n_{\text{i}}}{2}\right)}(\tau) - R_j^{\left(\frac{n_{\text{i}}}{2}\right)}(t) \right| = 0 \tag{3.221}$$

uniformly for $t \in [0, 2\pi]$ (as can be shown via a modification of the analysis in [35, pp. 208, 209] that takes into account the different interval of integration), so the sequences $\left(\mathcal{A}^m_{n_{\text{i}}}\right)$ are collectively compact and pointwise convergent [35, Theorem 12.12]. □

Theorem 3.2.3. *The sequences of numerical integration operators*

$$\mathcal{A}^m_{n_i}\varphi(t) \equiv \frac{L}{2\pi}\frac{2\pi}{n_i} \sum_{\left\{j:\, |t-t_j| < \frac{2\pi A_{sp}}{L}\right\}} K^m_{sp2}(t, t_j)\varphi(t_j) \tag{3.222}$$

on $C^0_{per}(2\pi)$ which approximate

$$\mathcal{A}^m\varphi(t) \equiv \frac{L}{2\pi} \int_{t-\frac{2\pi A_{sp}}{L}}^{t+\frac{2\pi A_{sp}}{L}} K^m_{sp2}(t,\tau)\varphi(\tau)\, d\tau \tag{3.223}$$

for $\varphi \in C^0_{per}(2\pi)$ are collectively compact and pointwise convergent.

Proof. As with $K^m_{\text{sp1}}(t,\tau)$ in Theorem 3.2.2, $K^m_{\text{sp2}}(t,\tau)$ can be periodically extended in τ (with period 4π) for each $t \in [0, 2\pi]$. Thus, applying the trapezoidal rule (2.136)—with an appropriately modified interval of integration—to $\mathcal{A}^m\varphi(t)$ results in the quadratures $\mathcal{A}^m_{n_{\text{i}}}\varphi(t)$. Because the trapezoidal rule is convergent, the sequences $\left(\mathcal{A}^m_{n_{\text{i}}}\right)$ are collectively compact and pointwise convergent [35, Theorem 12.8]. □

Theorem 3.2.4. *The sequences of numerical integration operators*

$$\mathcal{A}_{n_i}^m \varphi(t) \equiv \frac{L}{2\pi} \frac{2\pi}{n_i} \sum_{\left\{j:\ \frac{2\pi c_{sp} A_{sp}}{L} < |t-t_j| < \frac{2\pi A}{L}\right\}} K_{rp}^m(t,t_j)\varphi(t_j) \tag{3.224}$$

on $C_{per}^0(2\pi)$ *which approximate*

$$\mathcal{A}^m \varphi(t) \equiv \frac{L}{2\pi} \int_{t-\frac{2\pi A}{L}}^{t-\frac{2\pi c_{sp} A_{sp}}{L}} K_{rp}^m(t,\tau)\varphi(\tau)\, d\tau + \frac{L}{2\pi} \int_{t+\frac{2\pi c_{sp} A_{sp}}{L}}^{t+\frac{2\pi A}{L}} K_{rp}^m(t,\tau)\varphi(\tau)\, d\tau \tag{3.225}$$

for $\varphi \in C_{per}^0(2\pi)$ *are collectively compact and pointwise convergent.*

Proof. $P_{\mathrm{rp}}\left[x(t), x(\tau), c_{\mathrm{sp}}, A_{\mathrm{sp}}, c, A\right]$ is 0 for $\tau \in \left[t - \frac{2\pi c_{\mathrm{sp}} A_{\mathrm{sp}}}{L}, t + \frac{2\pi c_{\mathrm{sp}} A_{\mathrm{sp}}}{L}\right]$ and decays smoothly to 0 within a subinterval of $(t - 2\pi n_{\mathrm{per}}, t + 2\pi n_{\mathrm{per}})$, where $n_{\mathrm{per}} \equiv \left\lceil \frac{A}{L} \right\rceil$. Thus, $K_{\mathrm{rp}}^m(t,\tau)$ can be periodically extended in τ (with period $4\pi n_{\mathrm{per}}$) for each $t \in [0, 2\pi]$, and the application of the trapezoidal rule (2.136)—with an appropriately modified interval of integration—to $\mathcal{A}^m\varphi(t)$ results in the quadratures $\mathcal{A}_{n_{\mathrm{i}}}^m\varphi(t)$. Again, because the trapezoidal rule is convergent, the sequences $\left(\mathcal{A}_{n_{\mathrm{i}}}^m\right)$ are collectively compact and pointwise convergent [35, Theorem 12.8]. □

Theorem 3.2.1 immediately follows from the above results.

3.2.1.2 Existence and Convergence of Numerical Solutions

The convergence behavior of the quadratures of (3.211) carries over to the numerical solutions $\bar{\mu}_m^{A,n_{\mathrm{i}}}(t)$. The proof of the existence and convergence of these solutions is similar to the proof of Theorem 3.1.4 for the existence of the solutions $\mu_m^A(x)$ of the approximating scattering equations (3.1) and their convergence to the solutions $\mu_m(x)$ of the exact scattering equations (3.5) as $A \to \infty$.

Remark 3.2.2. *We recall that the solutions* $\bar{\mu}_m^A(t) \equiv \mu_m^A\left[x(t)\right]$ *of* (3.210) *exist if* k *is not a Wood Anomaly value and if* A *is sufficiently large (Theorem 3.1.4).*

Theorem 3.2.5. *Assume that the solutions* $\bar{\mu}_m^A \in C_{per}^0(2\pi)$ *of the approximating scattering equations* (3.210) *exist. Then, the solutions* $\bar{\mu}_m^{A,n_i} \in C_{per}^0(2\pi)$ *of the numerical approximation equations* (3.211) *exist for sufficiently large* n_i*, and these solutions converge in* n_i*, uniformly on* $[0, 2\pi]$*, to* $\bar{\mu}_m^A(t)$ *as* $n_i \to \infty$.

Proof. Given the results of Section 3.1.1 and Theorem 3.1.4, the operators

$$\mathcal{A}^m\varphi(t) \equiv \frac{L}{2\pi}\int_{t-\frac{2\pi A}{L}}^{t+\frac{2\pi A}{L}} K^m(t,\tau)\varphi(\tau)\,d\tau \tag{3.226}$$

on $C^0_{\text{per}}(2\pi)$ are compact, and $I \pm \mathcal{A}^m$ have bounded inverses on $C^0_{\text{per}}(2\pi)$. Also, the sequences of numerical integration operators

$$\begin{aligned}\mathcal{A}^m_{n_\text{i}}\varphi(t) \equiv \frac{L}{2\pi} &\sum_{\left\{j:\,|t-t_j|<\frac{2\pi A_\text{sp}}{L}\right\}} \left\{R_j^{\left(\frac{n_\text{i}}{2}\right)}(t)K^m_{\text{sp1}}(t,t_j) + \frac{2\pi}{n_\text{i}}K^m_{\text{sp2}}(t,t_j)\right\}\varphi(t_j)\\ &+ \frac{L}{2\pi}\frac{2\pi}{n_\text{i}} \sum_{\left\{j:\,\frac{2\pi c_\text{sp}A_\text{sp}}{L}<|t-t_j|<\frac{2\pi A}{L}\right\}} K^m_{\text{rp}}(t,t_j)\varphi(t_j)\end{aligned} \tag{3.227}$$

on $C^0_{\text{per}}(2\pi)$ which approximate $\mathcal{A}^m\varphi(t)$ for $\varphi \in C^0_{\text{per}}(2\pi)$ are collectively compact and pointwise convergent (Theorem 3.2.1). Therefore, $I \pm \mathcal{A}^m_{n_\text{i}}$ have bounded inverses on $C^0_{\text{per}}(2\pi)$—i.e., the solutions $\bar{\mu}_m^{A,n_\text{i}} \in C^0_{\text{per}}(2\pi)$ exist—for sufficiently large n_i [35, Theorem 10.9].

By [35, Corollary 10.11], we have the error estimate

$$\left|\left|\bar{\mu}_m^A - \bar{\mu}_m^{A,n_\text{i}}\right|\right|_\infty \leq C\left|\left|\left(\mathcal{A}^m - \mathcal{A}^m_{n_\text{i}}\right)\bar{\mu}_m^A\right|\right|_\infty \tag{3.228}$$

for sufficiently large n_i and some constant C. Since the quadratures $\mathcal{A}^m_{n_\text{i}}\bar{\mu}_m^A(t)$ converge in n_i, uniformly on $[0,2\pi]$, to $\mathcal{A}^m\bar{\mu}_m^A(t)$ as $n_\text{i} \to \infty$ (Theorem 3.2.1), we conclude that the solutions $\bar{\mu}_m^{A,n_\text{i}}(t)$ converge in n_i, uniformly on $[0,2\pi]$, to $\bar{\mu}_m^A(t)$ as $n_\text{i} \to \infty$. □

As discussed earlier, Theorem 3.2.5 immediately leads to the conclusion that the values

$$\tilde{\mu}_\ell = \hat{\mu}_\ell = \bar{\mu}_m^{A,n_\text{i}}(t_\ell), \quad \ell = 0,1,2,\ldots,n_\text{t}-1 \tag{3.229}$$

which solve the linear systems (3.212) for $n_\text{i} = n_\text{t}$ uniformly converge to $\bar{\mu}_m^A(t_\ell)$, $\ell = 0,1,2,\ldots,n_\text{t}-1$ as $n_\text{t} \to \infty$.

3.2.1.3 Super-Algebraic Convergence for Smooth Gratings

We now show that the convergence of the quadratures and the convergence of the numerical solutions are super-algebraic if the scattering surface is C^∞. By way of preparation, we recall that if the solutions $\bar{\mu}_m^A(t) \equiv \mu_m^A[x(t)]$ of the approximating scattering equations (3.210)

exist, then they belong to $C^{\infty}_{\text{per}}(2\pi)$ for cases with smooth grating profiles (Theorem 3.1.8).

Remark 3.2.3. *A straightforward extension of the following discussion can be made for finitely differentiable scattering surfaces.*

We first prove the following theorem regarding the quadratures.

Theorem 3.2.6. *Assume that the grating profile $f[x(t)]$ belongs to $C^{\infty}_{per}(2\pi)$ and that the solutions $\bar{\mu}^A_m \in C^{\infty}_{per}(2\pi)$ of the approximating scattering equations (3.210) exist. Then, the quadratures*

$$\begin{aligned}\mathcal{A}^m_{n_i}\bar{\mu}^A_m(t) \equiv \frac{L}{2\pi} \sum_{\left\{j:\, |t-t_j|<\frac{2\pi A_{sp}}{L}\right\}} \left\{ R_j^{\left(\frac{n_i}{2}\right)}(t) K^m_{sp1}(t,t_j) + \frac{2\pi}{n_i} K^m_{sp2}(t,t_j) \right\} \bar{\mu}^A_m(t_j) \\ + \frac{L}{2\pi}\frac{2\pi}{n_i} \sum_{\left\{j:\, \frac{2\pi c_{sp} A_{sp}}{L}<|t-t_j|<\frac{2\pi A}{L}\right\}} K^m_{rp}(t,t_j)\bar{\mu}^A_m(t_j)\end{aligned} \tag{3.230}$$

converge super-algebraically in n_i, uniformly on $[0,2\pi]$, to

$$\mathcal{A}^m\bar{\mu}^A_m(t) \equiv \frac{L}{2\pi}\int_{t-\frac{2\pi A}{L}}^{t+\frac{2\pi A}{L}} K^m(t,\tau)\bar{\mu}^A_m(t_j)\, d\tau \tag{3.231}$$

as $n_i \to \infty$.

Proof. As described in the proofs of Theorems 3.2.2–3.2.4, each of the kernel functions $K^m_{\text{sp1}}(t,\tau)$, $K^m_{\text{sp2}}(t,\tau)$ and $K^m_{\text{rp}}(t,\tau)$ can be periodically extended in τ (with periods of certain integer multiples of 4π) for each $t \in [0,2\pi]$. The profile $f[x(t)]$ being $C^{\infty}_{\text{per}}(2\pi)$ implies that these periodic extensions are C^{∞} in τ. Thus, the quadratures $\mathcal{A}^m_{n_{\text{i}}}\bar{\mu}^A_m(t)$ converge super-algebraically in n_{i}, uniformly on $[0,2\pi]$, to $\mathcal{A}^m\bar{\mu}^A_m(t)$ as $n_{\text{i}} \to \infty$, since Theorem 3.2.1 holds, the trapezoidal rule (2.136)—with appropriately modified intervals of integration—integrates exactly the trigonometric interpolation polynomials for $K^m_{\text{sp2}}(t,t_j)\bar{\mu}^A_m(t_j)$ and $K^m_{\text{rp}}(t,t_j)\bar{\mu}^A_m(t_j)$ [35, p. 201] and the quadrature formula (2.137)—with an appropriately modified interval of integration—integrates exactly the trigonometric interpolation polynomials for the values $K^m_{\text{sp1}}(t,t_j)\bar{\mu}^A_m(t_j)$ [35, p. 208]. □

Having this result, we now establish the super-algebraic convergence of the numerical solutions $\bar{\mu}^{A,n_{\text{i}}}_m(t)$.

Theorem 3.2.7. *Assume that the grating profile $f[x(t)]$ belongs to $C^{\infty}_{per}(2\pi)$, that the solutions $\bar{\mu}^{A}_{m} \in C^{\infty}_{per}(2\pi)$ of the approximating scattering equations (3.210) exist and that the solutions $\bar{\mu}^{A,n_i}_{m} \in C^{0}_{per}(2\pi)$ of the numerical approximation equations (3.211) exist. Then, the solutions $\bar{\mu}^{A,n_i}_{m}(t)$ are, in fact, $C^{\infty}_{per}(2\pi)$, and they converge super-algebraically in n_i, uniformly on $[0, 2\pi]$, to $\bar{\mu}^{A}_{m}(t)$ as $n_i \to \infty$.*

Proof. The solutions $\bar{\mu}^{A,n_{\mathrm{i}}}_{m}(t)$ are $C^{\infty}_{\mathrm{per}}(2\pi)$, because the functions $R_j^{\left(\frac{n_{\mathrm{i}}}{2}\right)}(t)$, $K^{m}_{\mathrm{sp1}}(t,t_j)$, $K^{m}_{\mathrm{sp2}}(t,t_j)$, $K^{m}_{\mathrm{rp}}(t,t_j)$ and $q_m[x(t)]$ in (2.165) are $C^{\infty}_{\mathrm{per}}(2\pi)$. Both Theorem 3.2.5 and Theorem 3.2.6 hold. So, given the error estimate (3.228) and the super-algebraic convergence of the quadratures (3.230), we conclude that the solutions $\bar{\mu}^{A,n_{\mathrm{i}}}_{m}(t)$ converge super-algebraically in n_{i}, uniformly on $[0, 2\pi]$, to $\bar{\mu}^{A}_{m}(t)$ as $n_{\mathrm{i}} \to \infty$. □

3.2.2 $\mathcal{O}(1)$ Computational Times for Simple-Reflection Cases

In Section 2.2.3, we undertook an examination of the integral (2.64)—a quantity that is closely related to (but considerably simpler than) the integral operator in the exact scattering equations (2.51) and (2.52). Using the smooth windowing function $P_1(0, x', c, A)$ (2.54) to formulate the approximation (2.75) to this simple integral, we demonstrated that as $k_n = k[1 + \sin(\theta)] + \frac{2\pi n}{L}$ increases we may allow the integration window size A to decrease and maintain the accuracy of our approximation. Thus, we expect that there is a similar relationship between k and A in the computational accuracy of our full algorithm.

Additionally, in Section 2.3.2 we motivated the development of the unknown $\mu_2(x)$ (2.101) and its use for simple-reflection cases. Both our physical intuition as well as the paper [17]—in which the ansatz (2.95) was successfully used—indicated that $\mu_2(x)$ is "slowly oscillating" (i.e., has a bounded number of oscillations as k increases) in such cases. We then showed in Section 2.3.3 that it indeed is advantageous to solve the scattering equations (2.100) for $\mu_2(x)$ rather than the equations (2.90) for $\mu_1(x)$ (2.88) when considering problems containing only simple reflections; in particular, Cases 1 and 2 of Section 2.3.3.5 together illustrated the slowly oscillating behavior of $\mu_2(x)$ when no multiple reflections occur.

Having in view the asymptotic series that is the slowly oscillating factor in the ansatz (2.95) as well as the regularity result from Theorem 3.1.7, we assume that for problems with smooth grating profiles and scattering configurations giving rise only to simple reflections we may

write the solutions $\mu_2(x)$ of the exact scattering equations (3.5) as

$$\mu_2(x) = \mu_2(x,k) \sim \sum_{n=0}^{\infty} \frac{\zeta_n(x)}{(kL)^n} \tag{3.232}$$

as $k \to \infty$; here $\zeta_n \in C^{\infty}_{\text{per}}(L)$ and does not depend upon k, and we write $\mu_2(x) = \mu_2(x,k)$ and use inverse powers of the dimensionless quantity kL for the sake of clarity in our later analysis. A theoretical basis for this assumption can be found in [45]. We show that under these conditions we may fix the quantity kA as $k \to \infty$ (for a given scattering profile $f(x)$ and incidence angle θ) and maintain a desired level of computational accuracy (Section 3.2.2.1). This implies that in such cases there are upper bounds on both the number of target points and the number of quadrature points that are needed for solving the linear systems (2.170), and we demonstrate that our algorithm therefore requires $\mathcal{O}(1)$ total computational time as $k \to \infty$ (for a given scattering profile $f(x)$ and incidence angle θ) in order to compute $\mu_2(x,k)$ to a fixed level of accuracy (Section 3.2.2.2).

3.2.2.1 $\mathcal{O}(1)$ Approximation Errors

Earlier, we established the super-algebraic convergence in A of the approximating solutions $\mu_m^A(x,k)$ to the exact solutions $\mu_m(x,k)$ as $A \to \infty$ for cases with smooth gratings (Theorems 3.1.9 and 3.1.10). Using that analysis, we here demonstrate that as k increases (for a given scattering profile $f \in C^{\infty}_{\text{per}}(L)$ and incidence angle θ) we may let A decrease—keeping kA fixed—and maintain a desired level of accuracy in approximating $\mu_2(x,k)$ by $\mu_2^A(x,k)$ for problems with scattering configurations that do not give rise to multiple reflections.

Remark 3.2.4. *As with Theorem 3.1.9, the analysis here can be modified appropriately for configurations with finitely differentiable scattering surfaces.*

We first prove a lemma that will allow us to easily demonstrate this error property given the ansatz (3.232) for $\mu_2(x,k)$.

Lemma 3.2.1. *Let the grating profile $f(x)$ belong to $C^{\infty}_{per}(L)$, and let $\varphi(x)$ belong to $C^{\infty}_{per}(L)$ as well. Also, let η be a real number such that $\eta > 0$. Given the operators K_A^2 (3.2) and K^2 (3.3), if for every $n \in \mathbb{Z}$*

$$|kL\,[1 \pm \sin(\theta)] + 2\pi n| > \eta > 0 \tag{3.233}$$

(i.e., if k is some distance away from all Wood Anomaly values), and if no multiple reflections are present, then

$$\max_{x\in[0,L]} \left|\left(K^2 - K_A^2\right)\varphi(x)\right| = \mathcal{O}\left((kL)^{-1}(kA)^{-p+\frac{1}{2}}\right) \tag{3.234}$$

for every integer $p \geq 1$ and every $A \in (0,\infty)$ as $k \to \infty$.

Proof. The proof of this lemma is closely related to the proof of Theorem 3.1.9. As such, we re-use much of the notation from that discussion. Recalling that

$$\left(K^2 - K_A^2\right)\varphi(x) = I_+^2(x,k,c,A) + I_-^2(x,k,c,A), \tag{3.235}$$

where $I_+^2(x,k,c,A)$ and $I_-^2(x,k,c,A)$ are given by (3.74) and (3.75), we only provide complete details for the estimation of the size of $I_+^2(x,k,c,A)$.

We employ the same changes of variables as in Section 3.1.3.1 in order to write

$$\begin{aligned} I_+^2(x,k,c,A) = \int_{\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty} \bar{P}_2\left(c, \frac{L\tilde{\phi}_+^{-1}}{cA}\right) \bar{g}_+\left(x,\tilde{\phi}_+^{-1}\right) \frac{h\left(kL\bar{u}_+\left(x,\tilde{\phi}_+^{-1}\right)\right)}{e^{ikL\bar{u}_+\left(x,\tilde{\phi}_+^{-1}\right)}} \\ \times e^{ikL[1+\sin(\theta)+c_1]\xi} e^{-ikLc_1\tilde{\phi}_+^{-1}} \bar{\varphi}\left[2\pi\left(\frac{x}{L} + \tilde{\phi}_+^{-1}\right)\right] \frac{\partial\tilde{\phi}_+^{-1}}{\partial\xi}\, d\xi, \end{aligned} \tag{3.236}$$

where $\tilde{\phi}_+^{-1} = \tilde{\phi}_+^{-1}(x,\xi)$ is the inverse of $\tilde{\phi}_+(x,t)$ (3.92) and where the lower limit of integration $\xi = \tilde{\phi}_+\left(x,\frac{cA}{L}\right)$ is the point at which $\bar{P}_2\left(c,\frac{L\tilde{\phi}_+^{-1}}{cA}\right)$ begins increasing from 0 as ξ increases. We may let $c_1 = 0$ since there are no multiple reflections (Remark 3.1.4), and this leads to the expression

$$\begin{aligned} I_+^2(x,k,c,A) = \int_{\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty} \bar{P}_2\left(c, \frac{L\tilde{\phi}_+^{-1}}{cA}\right) \bar{g}_+\left(x,\tilde{\phi}_+^{-1}\right) \frac{h\left(kL\bar{u}_+\left(x,\tilde{\phi}_+^{-1}\right)\right)}{e^{ikL\bar{u}_+\left(x,\tilde{\phi}_+^{-1}\right)}} \\ \times e^{ikL[1+\sin(\theta)]\xi} \bar{\varphi}\left[2\pi\left(\frac{x}{L} + \tilde{\phi}_+^{-1}\right)\right] \frac{\partial\tilde{\phi}_+^{-1}}{\partial\xi}\, d\xi. \end{aligned} \tag{3.237}$$

Thus, setting $\bar{\xi} \equiv kL\xi$, we have

$$\begin{aligned} I_+^2(x,k,c,A) = \frac{1}{kL}\int_{kL\tilde{\phi}_+\left(x,\frac{cA}{L}\right)}^{\infty} & \bar{P}_2\left[c, \frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right)\right]\bar{g}_+\left[x,\tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right)\right] \\ & \times \frac{h\left(kL\bar{u}_+\left[x,\tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right)\right]\right)}{e^{ikL\bar{u}_+\left[x,\tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right)\right]}} e^{i[1+\sin(\theta)]\bar{\xi}} \\ & \times \bar{\varphi}\left(2\pi\left[\frac{x}{L}+\tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right)\right]\right)\frac{\partial}{\partial\bar{\xi}}\tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right)kL\,d\bar{\xi}. \end{aligned} \tag{3.238}$$

The functions $\bar{P}_2\left[c, \frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right)\right]$, $e^{i[1+\sin(\theta)]\bar{\xi}}$ and $\bar{\varphi}\left(2\pi\left[\frac{x}{L}+\tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right)\right]\right)$ are $\mathcal{O}(1)$ in kL as $k\to\infty$. Expanding $\tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right)$ using the Taylor formula (3.99) and noting that $\tilde{\phi}_+^{-1}(x,0)=0$, we have

$$\begin{aligned} \tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right) = & \sum_{n=1}^{\gamma-1}\frac{1}{n!}\left(\frac{\bar{\xi}}{kL}\right)^n\left.\frac{\partial^n\tilde{\phi}_+^{-1}(x,z)}{\partial z^n}\right|_{z=0} \\ & + \frac{1}{(\gamma-1)!}\left(\frac{\bar{\xi}}{kL}\right)^{\gamma}\int_0^1(1-y)^{\gamma-1}\left.\frac{\partial^{\gamma}\tilde{\phi}_+^{-1}(x,z)}{\partial z^{\gamma}}\right|_{y\frac{\bar{\xi}}{kL}}dy \end{aligned} \tag{3.239}$$

for every integer $\gamma>1$. Additionally, it follows from (3.87) and (3.88) that both $\bar{g}_+(x,t)$ and $\bar{u}_+(x,t)$ admit similar expansions in t, with

$$\lim_{t\to 0^+}\bar{g}_+(x,t) = \frac{i}{4}\frac{\frac{h}{2L}(2\pi)^2\bar{f}''\left(\frac{2\pi x}{L}\right)}{1+\left(\frac{h}{2L}2\pi\right)^2\left[\bar{f}'\left(\frac{2\pi x}{L}\right)\right]^2} \tag{3.240}$$

and

$$\lim_{t\to 0^+}\frac{\bar{u}_+(x,t)}{t} = \sqrt{1+\left(\frac{h}{2L}2\pi\right)^2\left[\bar{f}'\left(\frac{2\pi x}{L}\right)\right]^2}. \tag{3.241}$$

Thus, $\frac{\partial}{\partial\bar{\xi}}\tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right)kL$, $\bar{g}_+\left[x,\tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right)\right]$ and $kL\bar{u}_+\left[x,\tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right)\right]$ also are $\mathcal{O}(1)$ in kL as $k\to\infty$. Finally, we note that

$$kL\tilde{\phi}_+\left(x,\frac{cA}{L}\right) = kcA\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right) \tag{3.242}$$

and

$$\frac{L}{cA}\tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right) = \frac{1}{kcA}kL\tilde{\phi}_+^{-1}\left(x,\frac{\bar{\xi}}{kL}\right); \tag{3.243}$$

here, $\frac{L}{cA}\tilde{\phi}_+\left(x,\frac{cA}{L}\right)$ (3.171) is $\mathcal{O}(1)$ in $\frac{A}{L}$ as $A\to\infty$, and—as can be easily deduced from the

Taylor expansion (3.239)—$kL\tilde{\phi}_+^{-1}\left(x, \frac{\bar{\xi}}{kL}\right)$ is $\mathcal{O}(1)$ in kL as $k \to \infty$. Since

$$\max_{x\in[0,L]} \left|I_+^2(x,k,c,A)\right| = \mathcal{O}\left(\left(\frac{A}{L}\right)^{-p+\frac{1}{2}}\right) \tag{3.244}$$

for every integer $p \geq 1$ as $A \to \infty$ (Theorem 3.1.9), the factor $\frac{1}{kL}$ in front of the integral (3.238), the factor kA in (3.242) and the factor $\frac{1}{kA}$ in (3.243) allow us to infer that

$$\max_{x\in[0,L]} \left|I_+^2(x,k,c,A)\right| = \mathcal{O}\left((kL)^{-1}(kA)^{-p+\frac{1}{2}}\right) \tag{3.245}$$

for every integer $p \geq 1$ and every $A \in (0,\infty)$ as $k \to \infty$.

The proof for $I_-^2(x,k,c,A)$—which can be written as (3.208) by employing certain changes of variables—follows analogously, since in this context we again may let $c_1 = 0$ (Remark 3.1.7) and since $|kL[1-\sin(\theta)] + 2\pi n| > \eta > 0$ for every $n \in \mathbb{Z}$. We conclude that

$$\begin{aligned}\max_{x\in[0,L]} \left|\left(K^2 - K_A^2\right)\varphi(x)\right| &\leq \max_{x\in[0,L]} \left|I_+^2(x,k,c,A)\right| + \max_{x\in[0,L]} \left|I_-^2(x,k,c,A)\right| \\ &= \mathcal{O}\left((kL)^{-1}(kA)^{-p+\frac{1}{2}}\right)\end{aligned} \tag{3.246}$$

for every integer $p \geq 1$ and every $A \in (0,\infty)$ as $k \to \infty$. □

Theorem 3.2.8. *Let the grating profile $f(x)$ belong to $C_{per}^{\infty}(L)$, and assume that the solutions $\mu_2(x,k)$ of the exact scattering equations (3.5) belong to $C_{per}^{\infty}(L)$ and can be written according to the ansatz (3.232). Also, let η be a real number such that $\eta > 0$. If for every $n \in \mathbb{Z}$*

$$|kL[1 \pm \sin(\theta)] + 2\pi n| > \eta > 0 \tag{3.247}$$

(i.e., if k is some distance away from all Wood Anomaly values), and if no multiple reflections are present, then the solutions $\mu_2^A(x,k)$ of the approximating scattering equations (3.1) exist for sufficiently large A and satisfy

$$\left|\left|\mu_2 - \mu_2^A\right|\right|_{\infty} \leq C(kL)^{-1}(kA)^{-p+\frac{1}{2}} \tag{3.248}$$

for every integer $p \geq 1$ and some constant C as $k \to \infty$.

Proof. Given the ansatz (3.232), it follows from Lemma 3.2.1 that

$$\max_{x\in[0,L]} \left|\left(K^2 - K_A^2\right)\mu_2(x,k)\right| = \mathcal{O}\left((kL)^{-1}\,(kA)^{-p+\frac{1}{2}}\right) \tag{3.249}$$

for every integer $p \geq 1$ and every $A \in (0,\infty)$ as $k \to \infty$. Applying Theorem 3.1.4, we conclude that the solutions $\mu_2^A(x,k)$ exist for sufficiently large A and that the inequality (3.248) holds for every integer $p \geq 1$ and some constant C as $k \to \infty$. □

Remark 3.2.5. *The above analysis can be modified straightforwardly to show that the exact integral operator K^2 has the property*

$$\max_{x\in[0,L]} \left|K^2\mu_2(x,k)\right| = \mathcal{O}\left(\frac{1}{kL}\right) \tag{3.250}$$

as $k \to \infty$ if k is some distance away from all Wood Anomaly values and if no multiple reflections are present, assuming that the ansatz (3.232) for $\mu_2(x,k)$ holds. In such cases, both $\mu_2(x,k)$ and $\mu_2^A(x,k)$ approach

$$\mu_2^0(x) = q_2(x) = \sin(\theta)f'(x) + \cos(\theta) \tag{3.251}$$

as $k \to \infty$ for a given scattering profile $f \in C_{per}^{\infty}(L)$ and incidence angle θ, where $q_2(x)$ is the right-hand side of (2.100) and $\mu_2^0(x)$ is the $A = 0$ solution of the approximating scattering equations (3.1). Equivalently, the highest order term $\zeta_0(x)$ of (3.232) satisfies

$$\zeta_0(x) = q_2(x); \tag{3.252}$$

this term corresponds to the Kirchhoff approximation just as the highest order term in the series of (2.95) corresponds to the Kirchhoff approximation (Section 2.3.2.1).

3.2.2.2 $\mathcal{O}(1)$ Numbers of Target and Quadrature Points

After proving a lemma that will help us establish the rate of necessary growth in the number n_{i} of integration points per period as k increases, we conclude with our main theorem—demonstrating the $\mathcal{O}(1)$ total computational time required as $k \to \infty$ (for a given scattering profile $f(x)$ and incidence angle θ) in order to compute $\mu_2(x,k)$ to a fixed level of accuracy if no multiple reflections are present.

Lemma 3.2.2. *The kernel $K^2(t,\tau)$ (2.142) in the approximating equations (3.210) has the property*

$$\max_{t\in[0,2\pi]} \left|K^2(t,\tau)\right| = \mathcal{O}\left(\sqrt{kA}\right) \tag{3.253}$$

as $k \to \infty$.

Proof. We have

$$\begin{aligned}\max_{t\in[0,2\pi]} \left|K^2(t,\tau)\right| &= \max_{t\in[0,2\pi]} \left|P_1\left[x(t),x(\tau),c,A\right] g\left[x(t),x(\tau)\right] \frac{h(ku\left[x(t),x(\tau)\right])}{e^{iku[x(t),x(\tau)]}} e^{ik\phi_m[x(t),x(\tau)]}\right| \\ &= \max_{t\in[0,2\pi]} \left|P_1\left[x(t),x(\tau),c,A\right] g\left[x(t),x(\tau)\right] h(ku\left[x(t),x(\tau)\right])\right|.\end{aligned} \tag{3.254}$$

We recall from (2.53) and (2.54) that $P_1\left[x(t),x(\tau),c,A\right] = 0$ for $|x(\tau)-x(t)| \geq A$ and that $P_1\left[x(t),x(\tau),c,A\right] = 1$ for $|x(\tau)-x(t)| \leq cA$. The estimate (3.253), therefore, follows from the application of the asymptotic formula (2.62) to the function

$$\begin{aligned}h(ku\left[x(t),x(\tau)\right]) &= ku\left[x(t),x(\tau)\right] H_1^1(ku\left[x(t),x(\tau)\right]) \\ &= k\sqrt{\left[x(t)-x(\tau)\right]^2 + \left\{f[x(t)]-f[x(\tau)]\right\}^2} \\ &\quad \times H_1^1\left(k\sqrt{\left[x(t)-x(\tau)\right]^2 + \left\{f[x(t)]-f[x(\tau)]\right\}^2}\right)\end{aligned} \tag{3.255}$$

as $k \to \infty$. □

Theorem 3.2.9. *Let the grating profile $f(x)$ belong to $C^\infty_{per}(L)$, and assume that the solutions $\mu_2(x,k)$ of the exact scattering equations (3.5) belong to $C^\infty_{per}(L)$ and can be written according to the ansatz (3.232). Also, let η be a real number such that $\eta > 0$. If for every $n \in \mathbb{Z}$*

$$|kL\left[1 \pm \sin(\theta)\right] + 2\pi n| > \eta > 0 \tag{3.256}$$

(i.e., if k is some distance away from all Wood Anomaly values), and if no multiple reflections are present, then

1. *$\mu_2(x,k)$ can be represented uniformly for $x \in [0,L]$ to a prescribed level of accuracy using an $\mathcal{O}(1)$ number of Fourier modes as $k \to \infty$;*

2. *$\mu_2(x,k)$ can be computed numerically to a fixed level of accuracy by solving the approximating linear systems (2.170) using an $\mathcal{O}(1)$ number of non-zero values $\hat{a}_{\ell,j}$ (2.167)*

as $k \to \infty$*;*

3. $\mu_2(x,k)$ *can be computed numerically to a fixed level of accuracy in* $\mathcal{O}(1)$ *total computational time as* $k \to \infty$*.*

the algorithm of this thesis—by solving the approximating linear systems (2.170)*—computes the solutions* $\mu_2(x,k)$ *at a prescribed level of numerical accuracy in* $\mathcal{O}(1)$ *computational time as* $k \to \infty$ *(for a given scattering profile* $f(x)$ *and incidence angle* θ*).*

Proof. We consider each part of the theorem in turn.

Proof of 1: Given the ansatz (3.232), for every $\epsilon > 0$ and integer $N \geq 0$ there exists a real $k_{\epsilon,N}$ such that

$$\max_{x\in[0,L]} \left| \mu_2(x,k) - \sum_{n=0}^{N} \frac{\zeta_n(x)}{(kL)^n} \right| < \frac{\epsilon}{2} \tag{3.257}$$

for $k > k_{\epsilon,N}$. Since each of the functions $\zeta_n(x)$ is independent of k, it follows that there exists an integer M (dependent upon $k_{\epsilon,N}, \epsilon, N$) such that

$$\max_{x\in[0,L]} \left| \sum_{n=0}^{N} \frac{\zeta_n(x)}{(kL)^n} - \sum_{m=-M}^{M} c_m(kL) e^{i\frac{2\pi m}{L}x} \right| < \frac{\epsilon}{2} \tag{3.258}$$

for $k > k_{\epsilon,N}$, where $c_m(kL)$ is the mth Fourier coefficient for the series $\sum_{n=0}^{N} \frac{\zeta_n(x)}{(kL)^n}$. Therefore,

$$\max_{x\in[0,L]} \left| \mu_2(x,k) - \sum_{m=-M}^{M} c_m(kL) e^{i\frac{2\pi m}{L}x} \right| < \epsilon \tag{3.259}$$

for $k > k_{\epsilon,N}$, i.e., $\mu_2(x,k)$ can be computed to any desired degree of precision using an $\mathcal{O}(1)$ number of Fourier modes as $k \to \infty$.

Proof of 2: In view of Part 1, our numerical method requires an $\mathcal{O}(1)$ number $n_{\rm t}$ of target points as $k \to \infty$ in order to compute $\mu_2(x,k)$ to a prescribed level of accuracy by solving the approximating linear systems (2.170).

In solving these linear systems, we may keep kA fixed as k increases without bound and ensure a desired degree of numerical precision for our solution (Theorem 3.2.8). Now, the number $n_{\rm i}$ of integration points per period must increase with k in order to maintain computational accuracy, since the number of oscillations of the kernel in the exact scattering equations (2.100)—and thus of the kernel $K^2(t,\tau)$ (2.142) in the approximating

equations (3.210)—increases with k. Given Lemma 3.2.2, this increase is linear in k for fixed kA because of the phase factor

$$e^{ik\phi_2[x(t),x(\tau)]} = e^{ik\left[\sqrt{[x(t)-x(\tau)]^2+\{f[x(t)]-f[x(\tau)]\}^2}-(\sin(\theta),-\cos(\theta))\cdot(x(t)-x(\tau),f[x(t)]-f[x(\tau)])\right]} \tag{3.260}$$

in (2.142). Therefore, we may leave $n_{\text{hwsp}} = \left\lfloor \frac{n_{\text{i}} A_{\text{sp}}}{L} \right\rfloor$ and $n_{\text{hw}} = \left\lfloor \frac{n_{\text{i}} A}{L} \right\rfloor$ constant as $k \to \infty$ and maintain computational accuracy, which implies that both the number of quadrature weights $R_j^{\left(\frac{n_{\text{i}}}{2}\right)}$ (2.168) and the number of integration points that lie within the integration window about each target point are $\mathcal{O}(1)$ as $k \to \infty$.

Thus, our method requires the computation of an $\mathcal{O}(1)$ number of values $\hat{a}_{\ell,j}$ (2.167) as $k \to \infty$.

Proof of 3: Each of the quadrature weights requires $\mathcal{O}(1)$ computational time as $n_{\text{i}} \to \infty$ in order to be computed (Section 2.5.3.1), which implies that each $\hat{a}_{\ell,j}$ of (2.170) is calculated in $\mathcal{O}(1)$ time. Additionally, as described in Section 2.5.3.2, we have an $\mathcal{O}(1)$ approach for determining the Fourier interpolation values $\hat{\mu}_j$ of (2.170) at the relevant integration points (the points that lie within the integration window about each target point). Therefore, we conclude that the algorithm of this thesis computes the solutions $\mu_2(x,k)$ at a prescribed level of numerical accuracy in $\mathcal{O}(1)$ computational time as $k \to \infty$. □

Chapter 4

Numerical Results

The numerical method described in this thesis was implemented in FORTRAN 77. "Double precision" and "double complex" data types were used as appropriate for machine-level-accurate computations. All of the numerical results we present were obtained from runs of this code on a 3.0 GHz Intel Xeon processor (2 MB cache).

For purposes of code verification, we computed scattering efficiencies for a variety of cases and compared these values to those generated by proven codes. The scattering configurations considered included cases from our two main parameter regimes—cases with multiple reflections (which may also include shadowing) and those with only simple reflections—so as to test the code's computations under the $\mu_1(x)$ and $\mu_2(x)$ representations of the density in their most appropriate settings (Sections 2.3.1 and 2.3.2). We also varied between TE/sound-soft and TM/sound-hard scattering, different values of k and different grating profiles. For those cases where k was not a Wood Anomaly value, the solver from [13]—previously mentioned in Section 2.2.1—was used as a baseline. Not only did both codes' computed efficiencies satisfy the energy balance criterion (2.45) to machine precision, but they also agreed with each other on an efficiency-by-efficiency basis. Additionally, previously published [17] efficiencies of certain cases which had Wood Anomaly values for k were reproduced. See Appendix B for examples.

Using this verified code, we undertook a convergence study to demonstrate the rapid convergence of our method in both number of discretization points and integration window size given a smooth grating profile (Section 4.1). The computational cost and accuracy of our method was compared against that of other rigorous methods (we restrict ourselves to comparing the performance of our solver to recent integral equation-based approaches that are the most efficient and accurate methods we have found in the literature) as well as

the non-convergent Kirchhoff approximation, using previously published results and, in the case of the solver from [13], running the code ourselves (Section 4.2); the code compares very favorably under a variety of scattering configurations, including those which simulate the real-life problem of scattering from an ocean surface. Additionally, for a grating of the form $f(x) = \frac{h}{2}\cos(2\pi x)$, we varied the height h as well as the incident wave's wavenumber k and incidence angle θ in order to see how the computational cost of our method varies with these parameters, including examining the sensitivity of our method for k at and near Wood Anomaly values (Section 4.3). The results of Section 4.3 further illuminate the capabilities of our method and demonstrate that its computational cost varies in accordance with the proofs given in earlier chapters of this thesis.

Remark 4.0.6. *The cases described in all of the sections in this chapter involve TE/sound-soft scattering. A selection of these cases is re-examined in Appendix C, where for each example the particular grating profile, wavenumber, incidence angle, representation of the "density" and set of numerical parameters are left unchanged but the type of scattering considered is TM/sound-hard instead of TE/sound-soft. In that study, we demonstrate that—all other things being equal—the type of scattering that is occurring does not significantly impact the accuracy of our solver.*

4.1 Convergence

In this section, we show that our numerical method yields rapidly convergent results for three typical scattering configurations—two which give rise to multiple reflections (one of which also contains shadowing) and one with only simple reflections. We do this by evaluating two types of quantities that are based upon the computed scattering efficiencies. First, since the scattering efficiencies e_n satisfy

$$\sum_{n\in U} e_n = 1, \tag{4.1}$$

which is the energy balance criterion (2.45) discussed in Section 2.1.4, one measure of the accuracy of a numerical solution is to calculate the error $\left|\sum_{n\in U} e_n - 1\right|$ for that solution's computed efficiencies; we call this error the "energy balance error." Second, for each configuration we use the method of [13] to compute a reference solution which has a very small

energy balance error (e.g., 10^{-13} to 10^{-16}); we evaluate the differences between a solution's computed efficiencies and those of the reference solution on an efficiency-by-efficiency basis, and we call the maximum of the absolute values of these differences the "maximum absolute error" ("max. abs. error"). We note that the energy balance error may continue decreasing without the maximum absolute error decreasing (see, e.g., Section 4.1.2). But, our method demonstrates convergence according to both of these measurements, i.e., in the sums of the scattering efficiencies and in each of the efficiencies as compared to those generated by another solver.

Remark 4.1.1. *See Appendix D for a brief discussion about rounding errors.*

4.1.1 Multiple-Reflection Cases

4.1.1.1 No Shadowing

We begin our convergence study by considering a grating profile of the form $f(x) = \frac{1}{2}\cos(2\pi x)$ and the incidence angle $\theta = 10°$, so that the scattering from this configuration includes multiple reflections but not shadowing (Figure 2.10). We take $\frac{k}{2\pi} = 10$, since this corresponds to an incident wave with a moderately sized wavenumber that is well away from the set of Wood Anomaly values. Using the solver described in [13], we generate a reference solution of this scattering problem with which we compare the solutions of our method; the reference solution has an energy balance error of 6.9×10^{-15} (see Figure 4.1 for a plot of its efficiencies).

Solving for $\mu_1(x)$, we first fix the discretization of the system, setting both the number n_t of target points per period and the number n_i of integration points per period to be 192, and we increase the integration window size A while keeping $A_{sp} = \frac{7}{8}$ (see Section 2.5 for a further description of these parameters). We check both the energy balance errors and the maximum absolute errors (relative to the reference solution) associated with the computed efficiencies. Confirming our analysis in Section 3.1, the efficiencies exhibit super-algebraic convergence in A up to the very small error level implicit in this discretization (Table 4.1). Similarly, super-algebraic convergence in n_t to the error level implicit for $A = 800$ is also achieved (as expected for our spectral method; see Section 3.2) when we fix $A = 800$ and increase n_t while keeping $n_i = n_t$ (Table 4.2). We also see rapid convergence in n_i when doubling it while leaving $A = 800$ and n_t fixed (Table 4.3), although we note that in this

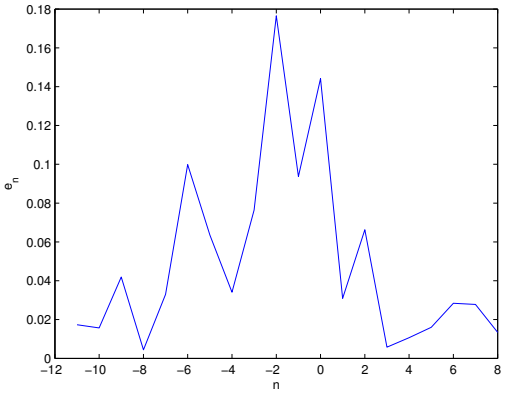


Figure 4.1: Efficiencies of the Regime 1 case with no shadowing

A	energy balance error	max. abs. error
100	8.8×10^{-6}	3.3×10^{-6}
200	1.5×10^{-7}	4.1×10^{-8}
400	8.0×10^{-11}	3.1×10^{-11}
600	1.8×10^{-12}	4.5×10^{-13}
800	8.5×10^{-13}	3.1×10^{-13}

Table 4.1: Convergence table for various A ($n_\mathrm{t} = n_\mathrm{i} = 192$) for the Regime 1 case with no shadowing

case further increases in n_i relative to each n_t (e.g., setting $n_\mathrm{i} = 3 \times n_\mathrm{t}$) yield no additional accuracy.

Remark 4.1.2. *Computing this case using* $n_t = 128$, $n_i = 128 \times 2$, $A = 800$ *results in an energy balance error of* 1.0×10^{-12} *and a maximum absolute error of* 3.0×10^{-13}. *This solution took* 56 *seconds to compute.*

n_t	energy balance error	max. abs. error
64	1.3×10^{0}	2.9×10^{-1}
96	2.3×10^{-4}	1.4×10^{-4}
128	2.0×10^{-6}	5.1×10^{-7}
160	2.1×10^{-12}	1.2×10^{-12}
192	8.5×10^{-13}	3.1×10^{-13}

Table 4.2: Convergence table for various n_t ($n_\mathrm{t} = n_\mathrm{i}$ and $A = 800$) for the Regime 1 case with no shadowing

n_t	n_i	energy balance error	max. abs. error
64	64×1	1.3×10^{0}	2.9×10^{-1}
64	64×2	3.1×10^{-1}	7.7×10^{-2}
96	96×1	2.3×10^{-4}	1.4×10^{-4}
96	96×2	1.4×10^{-11}	1.5×10^{-11}
128	128×1	2.0×10^{-6}	5.1×10^{-7}
128	128×2	1.0×10^{-12}	3.0×10^{-13}

Table 4.3: Convergence table for various n_t and n_i ($A = 800$) for the Regime 1 case with no shadowing

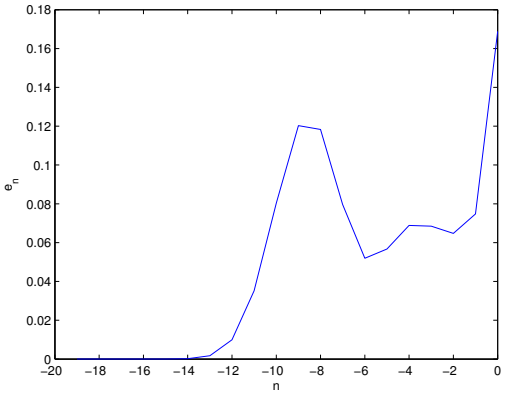


Figure 4.2: Efficiencies of the Regime 1 case with shadowing

4.1.1.2 Shadowing

We also examine scattering from the surface $f(x) = \frac{0.25}{2}\cos(2\pi x)$ by an incident wave with $\frac{k}{2\pi} = 10$ and $\theta = 75°$, since this configuration includes both multiple reflections and shadowing (Figure 2.10). The reference solution we use has an energy balance error of 1.7×10^{-15}, and its efficiencies are plotted in Figure 4.2. Solving for $\mu_1(x)$, our method yields convergence results similar to those in Section 4.1.1.1 (Tables 4.4–4.6). We note that for this case, unlike the previous one in Section 4.1.1.1, setting $n_\mathrm{i} = 3 \times n_\mathrm{t}$ does yield additional accuracy over the $n_\mathrm{i} = 2 \times n_\mathrm{t}$ solutions.

Remark 4.1.3. *The $n_t = 96$, $n_i = 96 \times 3$, $A = 750$ solution by our method was computed in 44 seconds.*

A	energy balance error	max. abs. error
100	5.5×10^{-7}	8.7×10^{-7}
200	5.5×10^{-9}	2.0×10^{-9}
400	2.8×10^{-12}	9.3×10^{-13}
600	1.4×10^{-14}	2.6×10^{-14}
750	1.7×10^{-15}	2.8×10^{-14}

Table 4.4: Convergence table for various A ($n_\mathrm{t} = 96,\ n_\mathrm{i} = 96 \times 3$) for the Regime 1 case with shadowing

n_t	energy balance error	max. abs. error
24	4.5×10^{-2}	2.4×10^{-2}
32	8.0×10^{-3}	5.4×10^{-3}
64	1.1×10^{-9}	8.1×10^{-10}
96	1.4×10^{-12}	1.2×10^{-12}
128	9.2×10^{-14}	9.2×10^{-14}

Table 4.5: Convergence table for various n_t ($n_\mathrm{i} = n_\mathrm{t}$ and $A = 750$) for the Regime 1 case with shadowing

n_t	n_i	energy balance error	max. abs. error
24	24×1	4.5×10^{-2}	2.4×10^{-2}
24	24×2	7.7×10^{-6}	3.3×10^{-6}
24	24×3	4.1×10^{-9}	4.8×10^{-8}
32	32×1	8.0×10^{-3}	5.4×10^{-3}
32	32×2	1.0×10^{-9}	7.7×10^{-10}
32	32×3	5.3×10^{-11}	6.5×10^{-11}
64	64×1	1.1×10^{-9}	8.1×10^{-10}
64	64×2	1.3×10^{-13}	1.1×10^{-13}
64	64×3	1.3×10^{-13}	5.1×10^{-14}
96	96×1	1.4×10^{-12}	1.2×10^{-12}
96	96×2	5.4×10^{-14}	3.6×10^{-14}
96	96×3	1.7×10^{-15}	2.8×10^{-14}

Table 4.6: Convergence table for various n_t and n_i ($A = 750$) for Regime 1 case with shadowing

4.1.2 Simple-Reflection Case

To consider a simple-reflection case, we let $f(x) = \frac{0.025}{2}\cos(2\pi x)$, $\frac{k}{2\pi} = 10$ and $\theta = 10°$. The reference solution has an energy balance error of 5.6×10^{-16}; Figure 4.3 contains a plot of the scattering efficiencies.

In view of the scattering phenomena which arise in this case, we use $\mu_2(x)$ for the unknown density, and the convergence results are similar to those of the cases considered previously. For our study of the convergence in A we consider two different discretizations: $n_{\mathrm{i}} = n_{\mathrm{t}} = 48$ and $n_{\mathrm{t}} = 16,\ n_{\mathrm{i}} = 16 \times 3$. Not only is super-algebraic convergence in A achieved up to the error levels implicit in these fixed discretizations, but also the errors for each discretization are nearly identical for most of the values of A considered (see Tables 4.7 and 4.8). Fixing $A = 30$, super-algebraic convergence in n_{t} (and n_{i} for $n_{\mathrm{i}} = n_{\mathrm{t}}$) is achieved up to machine precision (Table 4.9), while similarly rapid convergence in n_{i} (fixing $n_{\mathrm{i}} = 16$) is also demonstrated (Table 4.10).

We note that for $A = 20$ and $A = 30$ the individual efficiencies' errors for the $n_{\mathrm{t}} = 16,\ n_{\mathrm{i}} = 16 \times 3$ solutions do not decrease to machine precision (see Figure 4.4) as they do for the $n_{\mathrm{i}} = n_{\mathrm{t}} = 48$ solutions (see Table 4.7), and there is a similar stalling in convergence when fixing $A = 30$ and increasing the discretization from $n_{\mathrm{t}} = 16,\ n_{\mathrm{i}} = 16 \times 2$ to $n_{\mathrm{t}} = 16,\ n_{\mathrm{i}} = 16 \times 3$ (see Figure 4.4). This is due to the fact that not all of the significant Fourier modes of the density are being computed for $n_{\mathrm{t}} = 16$, while they are all computed for $n_{\mathrm{t}} = 48$ (see Figure 4.5). Nevertheless, very good results can be achieved for $n_{\mathrm{t}} = 16$ (often nearly identical to the $n_{\mathrm{t}} = 48$ results) with less computational time than is used for the $n_{\mathrm{t}} = 48$ solutions.

Remark 4.1.4. *The $n_t = n_i = 48,\ A = 30$ solution by our method was computed in* 0.24 *seconds, while the $n_t = 16,\ n_i = 16 \times 3,\ A = 30$ took* 0.09 *seconds. Another solution by our method has parameters $n_t = n_i = 38,\ A = 30$, took* 0.13 *seconds to compute, has an energy balance error of 4.4×10^{-16} and a maximum absolute error of 6.1×10^{-16}. We use this solution multiple times in Section 4.3.*

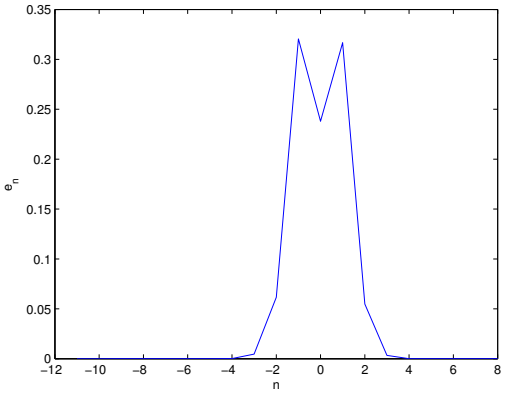


Figure 4.3: Efficiencies of the Regime 2 case

A	energy balance error	max. abs. error
1	6.6×10^{-5}	7.5×10^{-5}
2	2.8×10^{-6}	1.2×10^{-6}
5	3.0×10^{-8}	1.1×10^{-8}
10	9.6×10^{-11}	2.8×10^{-11}
20	2.8×10^{-14}	2.3×10^{-14}
30	8.9×10^{-16}	2.4×10^{-16}

Table 4.7: Convergence table for various A ($n_\mathrm{t} = n_\mathrm{i} = 48$) for the Regime 2 case

A	energy balance error	max. abs. error
1	6.6×10^{-5}	7.5×10^{-5}
2	2.8×10^{-6}	1.2×10^{-6}
5	3.0×10^{-8}	1.1×10^{-8}
10	9.6×10^{-11}	2.8×10^{-11}
20	3.0×10^{-14}	2.6×10^{-13}
30	1.1×10^{-15}	2.6×10^{-13}

Table 4.8: Convergence table for various A ($n_\mathrm{t} = 16$ and $n_\mathrm{i} = 16 \times 3$) for the Regime 2 case

n_t	energy balance error	max. abs. error
8	4.1×10^{0}	2.5×10^{0}
16	5.1×10^{-5}	1.8×10^{-5}
24	9.9×10^{-8}	6.9×10^{-8}
32	9.9×10^{-14}	1.4×10^{-13}
40	6.7×10^{-16}	9.4×10^{-16}
48	8.9×10^{-16}	2.4×10^{-16}

Table 4.9: Convergence table for various n_t ($n_\mathrm{i} = n_\mathrm{t}$ and $A = 30$) for the Regime 2 case

n_i	energy balance error	max. abs. error
16×1	5.1×10^{-5}	1.8×10^{-5}
16×2	1.0×10^{-13}	2.5×10^{-13}
16×3	1.1×10^{-15}	2.6×10^{-13}

Table 4.10: Convergence table for various n_i ($n_\mathrm{t} = 16$ and $A = 30$) for the Regime 2 case

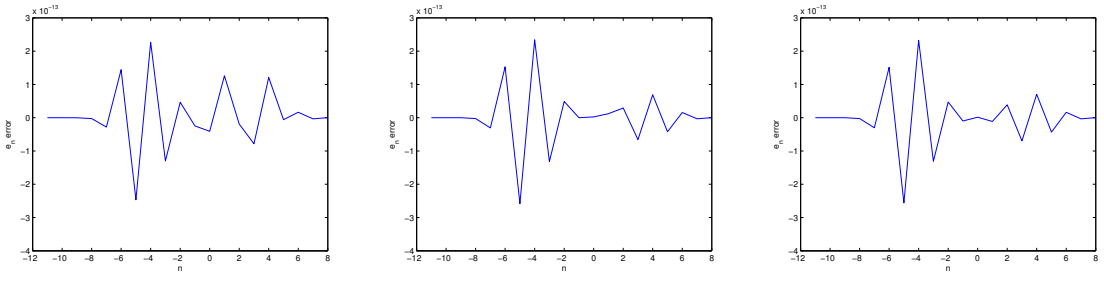


Figure 4.4: Errors of the $n_\mathrm{t} = 16,\ n_\mathrm{i} = 16 \times 2,\ A = 30$ solution (left), the $n_\mathrm{t} = 16,\ n_\mathrm{i} = 16 \times 3,\ A = 20$ solution (middle) and the $n_\mathrm{t} = 16,\ n_\mathrm{i} = 16 \times 3,\ A = 30$ solution (right) for the Regime 2 case

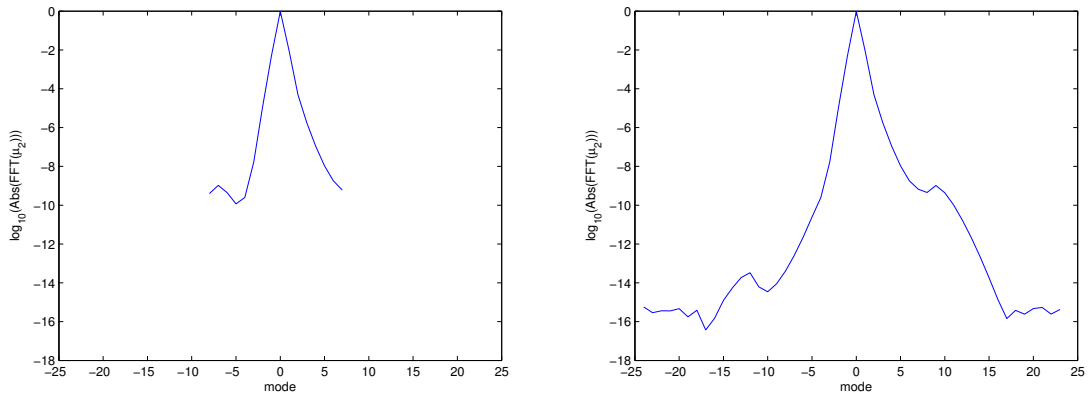


Figure 4.5: Fourier amplitudes for the $n_\mathrm{t} = 16,\ n_\mathrm{i} = 16 \times 3,\ A = 30$ solution (left) and the $n_\mathrm{t} = 48,\ n_\mathrm{i} = 48 \times 1,\ A = 30$ solution (right) for the Regime 2 case

4.2 Comparisons with Other Numerical Methods

We apply our numerical method to a variety of scattering configurations, including scattering from deterministic periodic gratings (Section 4.2.2) and from a randomly generated periodic rough surface which simulates ocean surface waves along one dimension (Section 4.2.3). To demonstrate the character of our algorithm, we compare its performance with those of some of the most efficient integral equation computational approaches available in the literature as well as with the approach based on the Kirchhoff approximation (KA). The results are satisfactory: the performance of our algorithm compares very favorably to those of other methods (in terms of both accuracy and efficiency) for a broad range of configurations.

4.2.1 Overview of Numerical Methods Used for Comparison

As mentioned above, as a basis of comparison we have used some of the most efficient algorithms available, including methods based on first-kind integral equations, second-kind integral equations, least-squares procedures and high-frequency approximations. A brief discussion of these approaches and our use of them in this thesis is provided in what follows.

4.2.1.1 Methods of [4]

A recently published paper [4] describes a first-kind integral equation formulation of the scattering problem and discusses three Galerkin methods—the "spectral-coordinate" (SC) and "spectral-spectral" (SS) methods that were previously presented in [21] as well as a modification of the SS method, called "SS*," that is introduced in this paper—along with a least-squares (LS) method that is not based on integral equations. For the Galerkin methods, the density is approximated for $x \in \left[-\frac{L}{2}, \frac{L}{2}\right]$ using a set of N basis functions (i.e., the number N is the number of degrees of freedom for the solution), while the LS method is derived using an N-term truncation of the Rayleigh expansion for the scattered field (see Section 2.1.2); each method uses its own set of basis functions, but each set of functions is closely related to the modes in the spectral expansion (2.22) of the periodic Green's function. Approximating linear systems of the form $Aa = b$ are developed for all four methods, where A denotes an $N \times N$ matrix corresponding to the kernel of the integral operator and b denotes a vector corresponding to the incident plane wave. Owing to the types of basis functions used in the Galerkin approaches, the vectors b in these cases can

be computed by means of simple function evaluations; the same is true of the matrix A in the case of the SC method (the functional expressions for the elements of this A come from approximating the full integral expressions of the elements by the midpoint rule). In contrast, the entries of A for the SS, SS* and LS methods (the entries are identical for the SS* and LS methods) and the elements of b for the LS method are integrals which are computed by M-point trapezoidal rule quadratures. Finally, we note that the SS* and LS methods are convergent, whereas the SC and SS methods may not be (as demonstrated by certain numerical results in [4]). Further details about all of these methods can be found in [4].

The comparisons we present in Section 4.2.2.1 include results generated by the application of these four algorithms. These results are taken from the work presented in [4].

Remark 4.2.1. *The paper [21]—in addition to presenting the SC and SS methods—discusses certain "coordinate-coordinate" (CC) methods denoted by "CC1," "CC2" and "CG." For the various scattering problems examined in that paper these CC methods seem to converge even when the SC and SS methods do not, but they are generally slower—sometimes orders of magnitude slower—than the SC and SS methods. Since the CC methods are so much slower than the SC and SS methods, and since they are not considered in [4] (a more recent study), we do not consider them here in this thesis.*

4.2.1.2 Method of [13]

The method of [13] is a collocation approach that computes n_{fl} modes in a truncation of the Floquet series expansion (2.131) for the unknown density in the second-kind integral equations (2.35) and (2.39) (TE/sound-soft scattering and TM/sound-hard scattering, respectively). The integral operator is approximated using an n_{ch}-term truncated Chebyshev series per Floquet mode. This results in n_{fl} equations in x, which are then discretized over the interval $x \in \left[-\frac{L}{2}, \frac{L}{2}\right]$ using n_{fl} uniformly spaced values of x so that an $n_{\text{fl}} \times n_{\text{fl}}$ system of equations results for the coefficients of the truncated Floquet series. An indication of the method used by these authors to produce the periodic Green's function and thus the kernel of the integral operator is given in Section 2.2.1; a key parameter required for the evaluation of this kernel, which was varied in our experiments to obtain optimal performance from this solver, is the number n_{pg} of points that are used in the Clenshaw-Curtis type quadrature for finite parts of the infinite integrals in these functions.

Remark 4.2.2. *The numbers denoted by n_{fl} and n_{ch} in this thesis are equal to the quantities $2N+1$ and $M+1$ of [13], respectively. The number n_{pg} is not directly described in [13], but it is a parameter in the code which we varied as needed for the present study. Both n_{ch} and n_{pg} are set in the code to be numbers of the form 2^n+1 (for integer n) for the sake of computational efficiency in computing certain FFTs.*

A number of comparisons between the results provided by our method and those resulting from the algorithm of [13] are given in Sections 4.2.2.1 and 4.2.2.2. The scattering configurations for the tests in the former section are drawn from [4], while those in the latter section concern multi-scale cases we designed for added generality in our test sets.

4.2.1.3 Method Using the Kirchhoff Approximation

The KA-based method uses the analytical approximation for the density given in (2.99); see Section 2.3.2.1 for details about its motivation. We discretize the analytical approximation using n_{ka} points, and we use a sufficiently large value of n_{ka} to produce the best accuracy possible for a given scattering case (as discussed earlier, this approximation of the density does not converge to the true density as $n_{ka} \to \infty$).

This approach is applied to all of the deterministic cases of Section 4.2.2 as well as the simulated ocean cases of Section 4.2.3.

4.2.2 Deterministic Grating Surfaces

The four numerical methods described in [4] are applied in that paper to three scattering configurations which also have been considered in the earlier paper [21]—configurations with deterministic periodic grating scattering surfaces and various incident fields. We apply our method, the method of [13] and the KA-based method to these same scattering problems, and we compare the results to each other and to those given in [4] (Section 4.2.2.1). Additionally, in Section 4.2.2.2 we consider scattering from a "multi-scale" surface which has a sinusoidal structure that is perturbed by a small but significantly more oscillatory component.

The performance of our algorithm compares favorably with those of these other methods over a variety of scattering configurations, including a broad range of wavenumbers, some configurations that only contain simple reflections, some that contain multiple reflections

(with or without shadowing), some where the scattering surface is a sinusoid and some where the scatterer is the multi-scale surface mentioned above. For example, for a given accuracy, our method requires many fewer degrees of freedom (unknowns) than those required by the methods of [4] for each configuration considered in that paper. Also, for some of the problems we discuss in Sections 4.2.2.1 and 4.2.2.2 the method of this thesis is significantly faster than that of [13] in computing solutions to any precision, while for others the method of [13] is faster or slower than our algorithm, depending upon whether full machine precision or less accurate solutions are required. Finally, while the very rapid KA-based method can be somewhat accurate for the particular high-frequency simple-reflection cases that we consider in this section, our method can compute the solutions for these cases much more accurately in short times, and it works well even when the KA-based approach breaks down—as it happens in the presence of multiple reflections.

4.2.2.1 Cases from [4]

The three scattering cases considered in [4] (and earlier in [21]) have grating surfaces of the form $f(x) = -\frac{h}{2}\cos\left(\frac{2\pi x}{L}\right)$. Besides the incidence angle θ, the physical parameters for these cases are given as the dimensionless quantities $\frac{h}{s\lambda}$ and $\frac{L}{\lambda}$, while the same cases are described in [21] using the dimensionless parameters $\frac{h}{L}$ and $\frac{\lambda}{L}$. Table 4.11 lists these parameters (using $L = 1$); Example 1 has simple reflections only, while Examples 2 and 3 have multiple reflections but no shadowing (Figure 2.10).

In [4], not only are the three Galerkin methods and the least-squares method applied to the scattering cases, but the Nyström method of [44] is also applied in order to generate reference solutions for purposes of comparison. The base 10 logarithm of the "energy balance error" (Section 4.1) and a second error measurement based upon the differences of the solutions' Rayleigh coefficients with those of the reference solutions (this measurement is the logarithm of something we call the "coefficients error" in later tables) are plotted in that paper. Also, the values of M used for the various trapezoidal rule quadratures of these methods are not given in [4], but it is stated that they are chosen sufficiently large so that the integrals are computed to machine accuracy. Additionally, the computational times for these cases are not given in the paper, but the values of N required, i.e., the numbers of degrees of freedom for the solutions, are emphasized there.

Taking the parameters as stated in Table 4.11, we compute the scattering efficiencies

for each of the three cases in [4] using our method, the method of [13] and the KA-based method. We choose numerical parameters for our solver and that of [13] so that both the energy balance errors and the maximum absolute errors (again, a reference solution using the method of [13] is computed for each case) of their solutions are near 10^{-16} (indicative of machine-level accuracies) as well as choosing other values for the parameters so that solutions with energy balance errors of approximately 1×10^{-4} (we call these "moderately accurate" or "mod. acc." solutions) are produced; for both sets of solutions the minimum values of these parameters necessary to achieve these error levels (as determined by extensive testing) are reported. We compare the energy balances and discretization levels to those given in [4] for the methods presented there. Additionally, we compare the computational times of our method and the method from [13], since the codes were run on the same computer and are thus directly comparable time-wise. To compare our method to the KA-based method, we choose a sufficient number of integration points for the KA-based approach so as to determine the maximum accuracy attainable by this non-convergent method.

Remark 4.2.3. *Computational timings for SS-generated and SC-generated solutions of these three scattering cases are given in [21]. For the first of these cases—the only one for which the SS and SC solutions are computed to machine precision in that paper—the SS method is much slower than the SC method due to the quadratures for the matrix elements for SS (the time to compute the matrix elements is denoted in [21] as the "fill time"). Specifically, for* $N = 128$ *the SS method had a fill time of* 4788 *seconds while the SC method had a fill time of* 0.64 *seconds; both methods required less than* 1 *second to compute the solution after the matrices were generated. But, it is unclear how the timings in [21] could be compared to those presented here, since that paper's results were generated using a much less recent computer (a SPARC 20 workstation). Therefore, we do not attempt to make such a comparison, but instead we compare the numbers of degrees of freedom of our solutions to the values of* N *reported in [4].*

One major result of this section is that the numbers of degrees of freedom for the solutions computed by the method of this thesis are significantly smaller than those for the solutions computed by the methods discussed in [4], with convergence being achieved for all three cases by the method of this thesis but not always being achieved by the SS

Example (in [4])	Example (in [21])	h	λ	θ
1	1A	0.075	0.01563553622559	20°
2	1B	0.075	0.01566499626662	75°
3	2A	0.25	0.95	20°

Table 4.11: Physical parameters for the cases that are described in both [4] and [21]

and SC methods. Our ability in these cases to represent the unknown "density" with a smaller number n_{t} of target points per period than the number n_{i} of integration points per period needed for our numerical quadrature gives the algorithm a strong advantage over the methods of [4] (which solve $N \times N$ linear systems, where N is the number of basis functions used to approximate the density). More precisely, the methods of [4] can achieve very accurate results by computing solutions with (in one case) as few as 2 (see Remark 4.2.4) "degrees of freedom per wavelength," where for a periodic scattering surface with arc length s over one period this number equals $\frac{N}{(s/\lambda)}$. The authors of [4] emphasize this fact since integral equation methods for these problems commonly require 5–10 degrees of freedom per wavelength (as is also stated in the survey paper [56] referred to in Section 1.1). Our method, however, requires significantly fewer degrees of freedom per wavelength to compute the solutions of these problems at (or near) machine precision, particularly when there is no multiple scattering.

Remark 4.2.4. *The values s/λ for the scattering cases are incorrectly doubled in [4], so that the numbers of degrees of freedom per wavelength for its methods are understated in that paper by a factor of 2. In accordance with the formula (2.41) for the differential arc length $ds(\boldsymbol{r}')$, we numerically evaluate*

$$s/\lambda = \frac{1}{\lambda} \int_{-\frac{L}{2}}^{\frac{L}{2}} \sqrt{1 + [f'(x')]^2}\, dx' \tag{4.2}$$

for each of the scattering cases later in this section.

Additionally, this thesis' approach performs well relative to the method introduced in [13]. It is dramatically faster than that algorithm in the simple-reflection case, and it takes a similar amount of time in the other cases. These comparisons help demonstrate the efficiency that arises from our solver's use of small values of the integration window size A for certain configurations or if less accuracy (such as that of the "mod. acc." solutions)

is required.

Finally, while the KA-based method is somewhat accurate in the simple-reflection case, it is not at all accurate in the others due to the nature of the approximation it uses. Our method, however, is accurate with short computational times for all three of the problems in this section.

Example 1: For this case $\lambda = 0.01563553622559 \longrightarrow \frac{k}{2\pi} \approx 63.95687$. Also, $s \approx 65\lambda$. We compute $\mu_2(x)$ when applying our method, since this case has no multiple reflections or shadowing. But, we also use our method to compute $\mu_1(x)$ in order to show this still leads to very fast results. The efficiencies of this scattering configuration (according to the reference solution) are plotted in Figure 4.6.

Remark 4.2.5. *The value for $\frac{k}{2\pi}$, i.e., for $\frac{L}{\lambda}$ with $L = 1$, is incorrectly stated as* 63.9587 *in [4]; this appears to be simply a typographical error. Additionally, in that paper the value for s is given as being approximately 130λ—a doubling of the correct value.*

The methods of [4] compute the solutions very accurately (with energy balance errors of 10^{-15} to 10^{-16}) with $N = 128$ degrees of freedom; see Table 4.12, which lists N as well as the energy balance errors and coefficients errors (these errors are estimated from plots in [4]). This corresponds to solutions that are very close to machine precision with slightly less than 2 (not 1 as stated in [4]) degrees of freedom per wavelength. As noted in [4], the coefficients errors are somewhat larger than the energy balance errors. This may indicate that the individual efficiencies computed by the four methods are somewhat less accurate than suggested by energy balance errors, although only the energy balance of the reference Nyström solution is given and it is unclear how accurate this reference solution's Rayleigh coefficients (and thus its efficiencies) are.

Our method, however, only requires $n_{\mathrm{t}} = 28$ target points per period to accurately compute $\mu_2(x)$ (appropriate for this scattering case) to machine precision (Table 4.13), which is only about $\frac{n_{\mathrm{t}}}{(s/\lambda)} = 0.43$ degrees of freedom per wavelength. This is accomplished because of our method's ability to have many fewer target points per period than the number n_{i} of integration points per period (for this solution $n_{\mathrm{i}} = 28 \times 10 = 280$) used to compute the integral operator. This is certainly an improvement over the 2 degrees of freedom per wavelength required by the methods of [4] as well as the general rule of thumb of 5–10 degrees of freedom per wavelength needed by other integral equation methods. Even if

we use our method to compute $\mu_1(x)$ for this case, only about 1.2 degrees of freedom per wavelength ($n_t = 80$) are required for machine precision.

Remark 4.2.6. *We also may infer that for this case our solver performs significantly better than the Nyström method of [44], since the authors of [4] note that the reference solution generated by that method comes from solving a linear system "over* 20 *times larger" than each of the systems used to generate the* $N = 128$ *solutions.*

Further comparisons with the performance offered by the methods described in [4] are difficult to make, as we stated previously. We do not know the values of M used to compute the quadratures by the methods of [4]. Also, no computational times are given in that paper to compare with our method's total time of 0.2 seconds for this case. As we noted in Remark 4.2.3, computational times for the SS and SC methods are given in [21] for this case (and the others) as computed on a SPARC 20 workstation, but precise comparisons time-wise cannot be made.

This thesis' approach is also significantly faster than the method of [13] in solving this scattering problem. In particular, it takes 0.2 seconds compared to that method's 37 seconds to achieve machine precision accuracy (see Tables 4.13 and 4.14): nearly 200 times faster. It is similarly faster (0.04 seconds vs. 8 seconds) in computing a "mod. acc." solution; Figure 4.6 includes plots of the efficiency errors of the "mod. acc." solutions, and on an efficiency-by-efficiency basis the solution produced by the approach of [13] is actually somewhat less accurate than that generated by our method, although the energy balance errors are very similar. For both the machine precision and "mod. acc." solutions our approach uses about one-third of the numbers of degrees of freedom that the method of [13] require.

Even if we solve for $\mu_1(x)$ by our method in order to compute the scattering efficiencies, the total computing times for both the machine precision and "mod. acc." data are significantly less than those for the approach of [13] (a factor of 200 smaller for the "mod. acc." data and just under a factor of 100 smaller for the machine precision data). The numbers of degrees of freedom of the solutions are very similar between the two methods, however. This should come as no surprise: the coefficients of the Fourier series of $\mu_1(x)$ equal the Floquet series amplitudes computed by the method of [13] multiplied by a constant factor (Section 2.3.3.5).

Method	N	energy balance error	coefficients error
SS*	128	$\approx 1 \times 10^{-16}$	$\approx 3 \times 10^{-13}$
SS	128	$\approx 1 \times 10^{-16}$	$\approx 3 \times 10^{-13}$
SC	128	$\approx 1 \times 10^{-15}$	$\approx 3 \times 10^{-13}$
LS	128	$\approx 3 \times 10^{-16}$	$\approx 3 \times 10^{-13}$
Nyström	N/A	$\approx 3 \times 10^{-15}$	—

Table 4.12: Results of the four methods of [4] plus the Nyström method of [44] for Example 1

rep.	n_t	n_i	A_{sp}	A	energy balance error	max. abs. error	time (sec)
$\mu_2(x)$	12	12×10	0.21875	0.25	1.4×10^{-4}	2.9×10^{-5}	0.04
$\mu_2(x)$	28	28×10	0.875	6	3.1×10^{-15}	2.6×10^{-15}	0.2
$\mu_1(x)$	38	38×3	0.21875	0.25	8.5×10^{-5}	4.8×10^{-4}	0.04
$\mu_1(x)$	80	80×3	0.875	6	1.8×10^{-15}	2.9×10^{-15}	0.4

Table 4.13: Results of this thesis' method for Example 1

The KA-based approach generates a solution with a slightly smaller energy balance error than that of our method's "mod. acc." solution (using $\mu_2(x)$) and does so in slightly less time (Tables 4.13 and 4.15), although the two solutions have very similar efficiency-by-efficiency errors (Figure 4.6). With very little additional computational time, however, machine-level-accurate data is obtained by our approach, while no additional accuracy is possible with the KA-based method.

Example 2: Here $\lambda = 0.01566499626662 \longrightarrow \frac{k}{2\pi} \approx 63.83659$. Although the value of λ is slightly different than that for Example 1, the relationship $s \approx 65\lambda$ still holds. For this case, however, we only compute $\mu_1(x)$ since this is a (no shadowing) multiple-reflection configuration. Figure 4.7 contains a plot of the scattering efficiencies.

Again, the numbers of degrees of freedom required for the solutions generated by the method of this thesis are much smaller than the numbers for those solutions of similar (or lesser) accuracy that were computed by the approaches of [4]. Table 4.16 lists the results

n_{fl}	n_{ch}	n_{pg}	energy balance error	max. abs. error	time (sec)
35	257	65	8.9×10^{-5}	2.5×10^{-3}	8
81	513	257	1.3×10^{-15}	3.9×10^{-16}	37
161	1025	513	1.1×10^{-16}	—	152

Table 4.14: Results of the method of [13] for Example 1

n_{ka}	energy balance error	max. abs. error	time (sec)
400	2.5×10^{-6}	8.1×10^{-5}	0.02

Table 4.15: Result of the KA-based method for Example 1

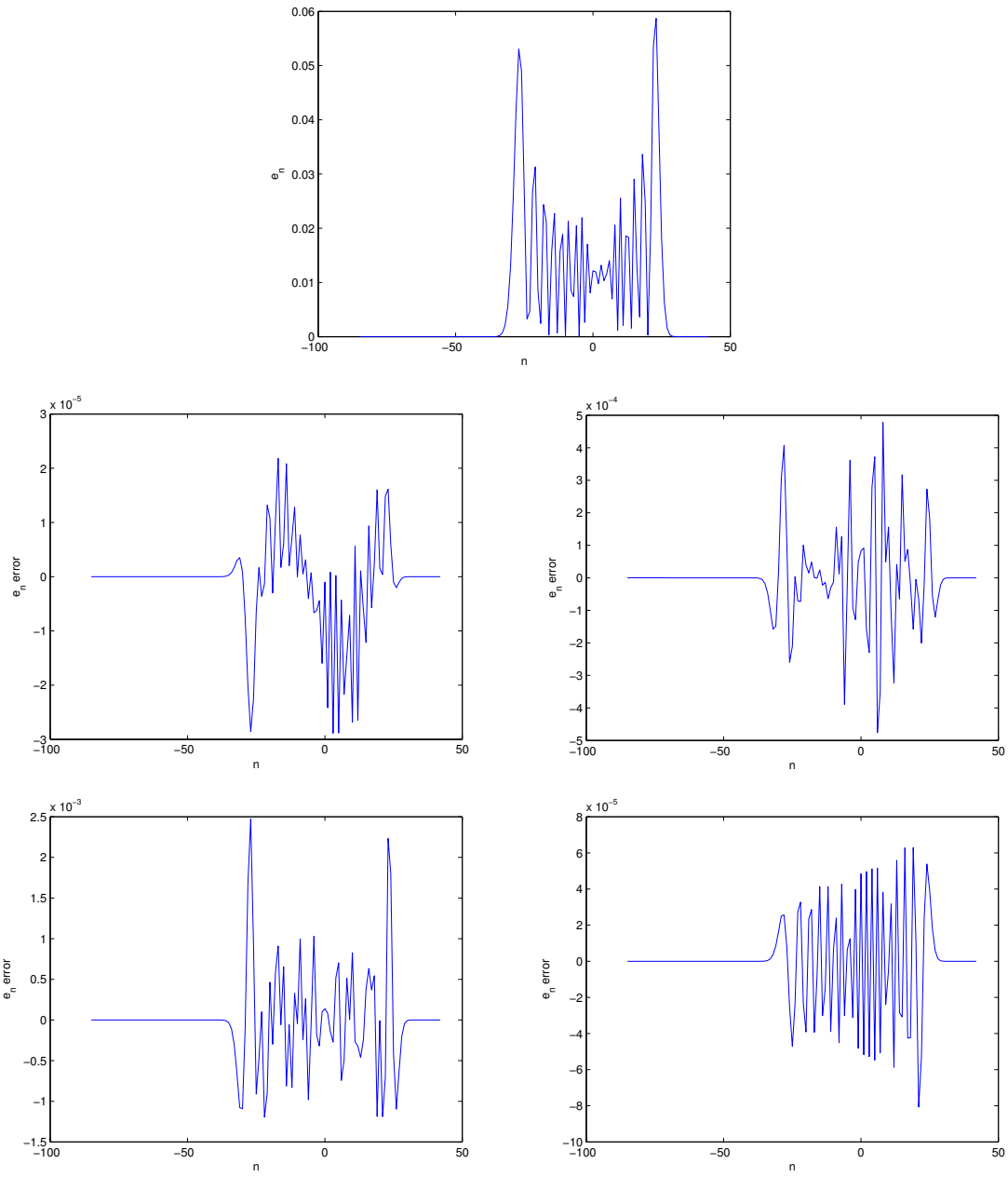


Figure 4.6: Efficiencies (top), errors of this work's "mod. acc." solutions ($\mu_2(x)$ on the middle-left and $\mu_1(x)$ on the middle-right), errors of the "mod. acc." solution produced by the method of [13] (bottom-left) and errors of the KA-based method's solution (bottom-right) for Example 1

from [4] for two different values of N—$N = 148$ and $N = 208$—which correspond to certain methods' efficiencies having energy balance errors of approximately 10^{-4} and the smallest listed in the paper for this example, respectively. Our method requires only $n_{\mathrm{t}} = 20$ target points to achieve error levels at least slightly better than those produced by the methods of [4] with $N = 148$; see Table 4.17. Also, near-machine-precision accuracy is obtained by our approach with $n_{\mathrm{t}} = 48$—the error levels only being matched by SS's $N = 208$ solution, while the other methods from [4] perform significantly worse for $N = 208$. In terms of number of degrees of freedom per wavelength, our method's more accurate solution has approximately 0.74—extremely good given its accuracy—while SS's $N = 208$ solution has approximately 3.2.

Our method also performs satisfactorily relative to the approach of [13], although in this case the performance is not as clearly superior as it is for Example 1. Since the thesis' approach employs $\mu_1(x)$ for the solution of this problem, the numbers of degrees of freedom for its solutions (given the error levels achieved) are nearly identical to those of the method of [13] (Tables 4.17 and 4.18; Figure 4.7 indicates that the individual efficiencies' errors are similar between the two methods' "mod. acc." solutions). The computational time used by our solver to compute its "mod. acc." solution is much smaller than the time needed to do the same by the method of [13] (0.7 seconds vs. 10 seconds). But, it takes 48 seconds to compute its more accurate solution, while the approach of [13] only requires 25 seconds to compute a similarly accurate solution.

The increase in each method's computational times can be explained in terms of each solver's numerical parameters. For both methods, the vast majority of the computational times for this case are spent in building the linear systems to be solved; the times needed to build such systems for the methods of [4] are called "fill times" (Remark 4.2.3). For our algorithm, this time increases linearly with n_{t}, n_{i} and A, while for the method of [13] it increases linearly in n_{fl} and n_{ch} but very slowly with respect to n_{pg}. The more accurate solution for our algorithm has increases in the values of these key parameters over the values for the "mod. acc." solution by factors of $\frac{48}{20} = 2.4$, $\frac{576}{200} = 2.88$ and $\frac{800}{70} \approx 11.4$. The product of these factors is approximately 79, which is very close to the ratio of times $\frac{48 \text{ seconds}}{0.7 \text{ seconds}} \approx 69$ (there is some overhead in both of these computational times which has not been factored in). The two solutions by the solver of [13] differ primarily in an increase in n_{fl} from 21 to 51 (n_{ch} is the same for these solutions), and this increase is matched by

Method	N	energy balance error	coefficients error
SS*	148	$\approx 1 \times 10^{-4}$	$\approx 3 \times 10^{-3}$
SS	148	$\approx 3 \times 10^{-1}$	$\approx 1 \times 10^{0}$
SC	148	$\approx 1 \times 10^{-3}$	$\approx 1 \times 10^{-2}$
LS	148	$\approx 3 \times 10^{-4}$	$\approx 3 \times 10^{-4}$
SS*	208	$\approx 1 \times 10^{-8}$	$\approx 1 \times 10^{-7}$
SS	208	$\approx 1 \times 10^{-13}$	$\approx 3 \times 10^{-11}$
SC	208	$\approx 1 \times 10^{-9}$	$\approx 1 \times 10^{-8}$
LS	208	$\approx 1 \times 10^{-9}$	$\approx 1 \times 10^{-9}$
Nyström	N/A	$\approx 3 \times 10^{-14}$	—

Table 4.16: Results of the four methods of [4] plus the Nyström method of [44] for Example 2

n_{t}	n_{i}	A_{sp}	A	energy balance error	max. abs. error	time (sec)
20	20×10	0.875	70	7.7×10^{-5}	6.1×10^{-5}	0.7
48	48×12	0.875	800	9.1×10^{-14}	9.1×10^{-14}	48

Table 4.17: Results of this thesis' method for Example 2

the increase in computational time from 10 seconds to 25 seconds.

The KA-based method generates a solution that is rather inaccurate (Table 4.19), with errors of the same order as the sizes of some of the largest scattering efficiencies (Figure 4.7). This is due to the presence of multiple reflections which are not accounted for in the approximation. Our method's "mod. acc." solution is much more accurate, and it is computed in only 0.7 seconds. Of course, our approach is able to generate even more accurate solutions in short computing times if desired, as shown by the $n_{\mathrm{t}} = 48$ solution, but the KA-based method cannot do the same.

Example 3: This configuration has $\lambda = 0.95 \longrightarrow \frac{k}{2\pi} \approx 1.05263$, and $s \approx 1.2\lambda$ (the grating here is deeper than in the earlier examples). Again, since there are multiple reflections (but no shadowing), we compute $\mu_1(x)$ for our algorithm. Our scattering efficiencies results are described using tables instead of plots since there are only two efficiencies for this case;

n_{fl}	n_{ch}	n_{pg}	energy balance error	max. abs. error	time (sec)
21	513	33	1.4×10^{-4}	3.5×10^{-5}	10
51	513	513	3.9×10^{-15}	5.0×10^{-16}	25
101	1025	1025	5.4×10^{-15}	—	108

Table 4.18: Results of the method of [13] for Example 2

n_{ka}	energy balance error	max. abs. error	time (sec)
400	1.6×10^{-1}	6.3×10^{-2}	0.02

Table 4.19: Result of the KA-based method for Example 2

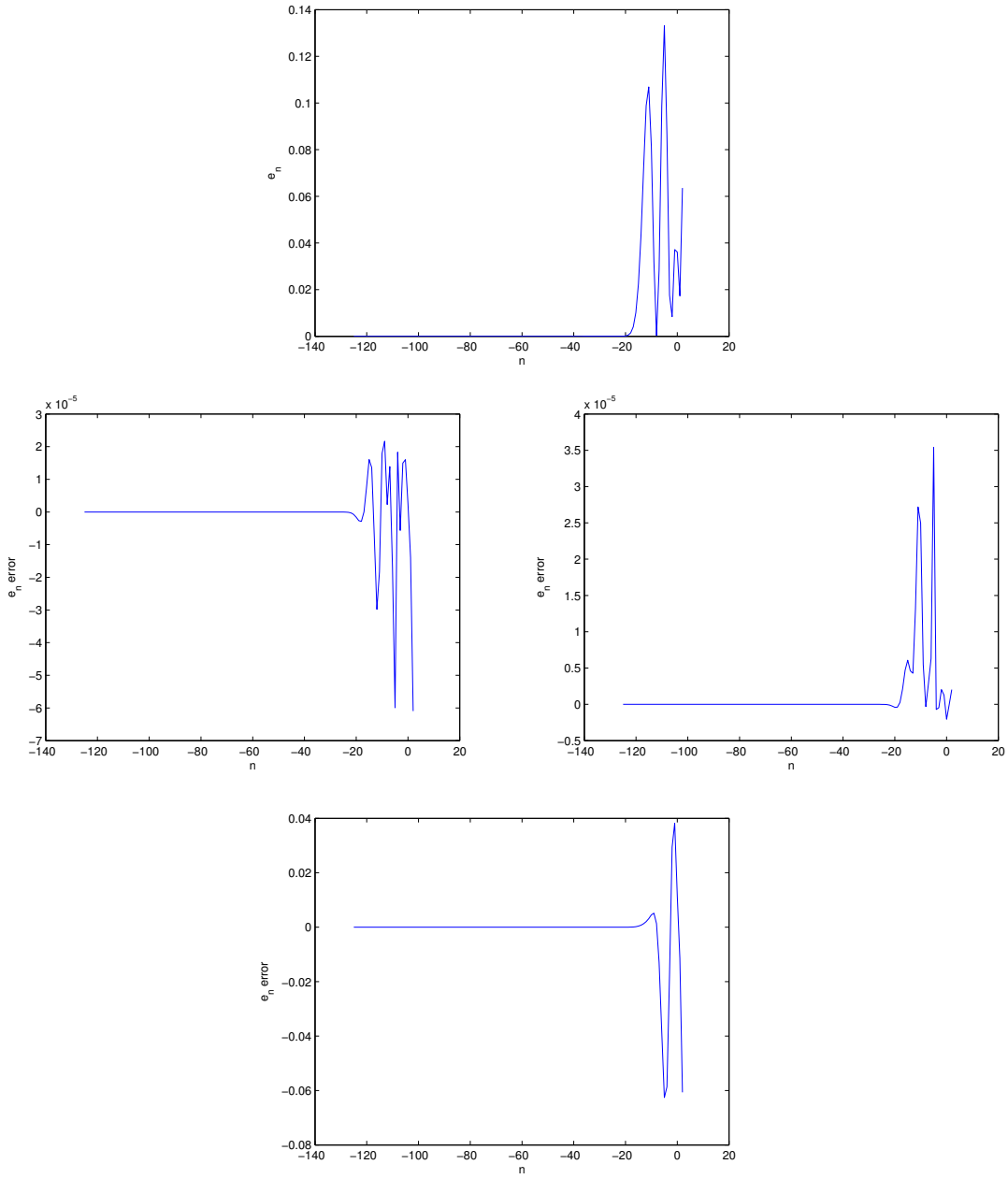


Figure 4.7: Efficiencies (top), errors of this work's "mod. acc." solution (middle-left), errors of the "4 digit" solution produced by the method of [13] (middle-right) and errors of the KA-based method's solution (bottom) for Example 2

Table 4.24 lists the efficiencies.

The SS and SC methods of [4] do not converge as the value of N is increased from 18 to 66 (Table 4.20). The SS* and LS methods do converge, however, with good accuracy for $N = 34$ (approximately 28 degrees of freedom per wavelength) and $N = 66$ (approximately 55 degrees of freedom per wavelength). The $N = 66$ solutions by these methods are the most accurate, with energy balance errors of approximately 10^{-9} in size.

Our method, though, not only converges for this case, but it also produces a solution of near-machine-precision accuracy with $n_t = 16$ target points—approximately 13 degrees of freedom per wavelength (Table 4.21). Also, accuracy levels similar to those of the approaches of [4] when using $N = 18$ are obtained by our solver for $n_t = 8$ (about 7 degrees of freedom per wavelength).

As in Example 2, the performance of our algorithm in this case is comparable to that of the method of [13] in that somewhat less time is needed for its "mod. acc." solution but somewhat more time is required for its more accurate solution (Tables 4.21 and 4.22; Tables 4.24 shows the close similarity between our method's "mod. acc." solution and that of the approach of [13]). In this case both solvers compute extremely accurate data in about 1 or 2 seconds, using many fewer degrees of freedom in their solutions than the methods of [4] use to compute solutions having at least 6 digits less accuracy.

The KA-based method again (like in Example 2) is very rapid but fails to compute the scattering efficiencies to a useful degree of accuracy (Tables 4.23 and 4.24) due to the presence of multiple reflections. So, our approach's "mod. acc." efficiencies (taking only 0.03 seconds to compute) are significantly more accurate, and even much more accurate efficiencies are obtainable by our method in short times.

4.2.2.2 Multi-Scale Cases

In addition to the cases from [4], we also examine scattering from the multi-scale surface $f(x) = \frac{0.025}{2}\left[\cos(2\pi x) + 0.04\sin(50\pi x)\right]$. The configurations discussed in this section have incident plane waves with incidence angles $\theta = 30°$ and $\theta = 85°$, and for each angle there are two cases: one with wavenumber k such that $\frac{k}{2\pi} = 10.5$ (a wavenumber in the "resonance" regime) and another with $\frac{k}{2\pi} = 1000.5$ (a wavenumber in the "high-frequency" regime). The wavenumbers are chosen so as to avoid Wood Anomalies. Also, the cases with $\frac{k}{2\pi} = 1000.5$ are similar to the simulated ocean surface problems in Section 4.2.3 in terms of the size and

Method	N	energy balance error	coefficients error
SS*	18	$\approx 3 \times 10^{-5}$	$\approx 1 \times 10^{-4}$
SS	18	$\approx 3 \times 10^{-3}$	$\approx 1 \times 10^{-2}$
SC	18	$\approx 1 \times 10^{-3}$	$\approx 1 \times 10^{-2}$
LS	18	$\approx 1 \times 10^{-4}$	$\approx 1 \times 10^{-4}$
SS*	34	$\approx 1 \times 10^{-7}$	$\approx 3 \times 10^{-7}$
SS	34	$\approx 1 \times 10^{-2}$	$\approx 6 \times 10^{-2}$
SC	34	$\approx 1 \times 10^{-2}$	$\approx 6 \times 10^{-2}$
LS	34	$\approx 3 \times 10^{-7}$	$\approx 3 \times 10^{-7}$
SS*	66	$\approx 3 \times 10^{-9}$	$\approx 1 \times 10^{-8}$
SS	66	$\approx 1 \times 10^{-1}$	$\approx 1 \times 10^{0}$
SC	66	$\approx 3 \times 10^{-1}$	$\approx 1 \times 10^{0}$
LS	66	$\approx 1 \times 10^{-9}$	$\approx 1 \times 10^{-8}$
Nyström	N/A	$\approx 1 \times 10^{-13}$	—

Table 4.20: Results of the four methods of [4] plus the Nyström method of [44] for Example 3

n_t	n_i	A_sp	A	energy balance error	max. abs. error	time (sec)
8	8×1	0.875	30	8.8×10^{-5}	1.3×10^{-4}	0.03
16	16×6	0.875	600	4.1×10^{-15}	1.0×10^{-14}	2.1

Table 4.21: Results of this thesis' method for Example 3

n_fl	n_ch	n_pg	energy balance error	max. abs. error	time (sec)
7	17	17	1.4×10^{-4}	1.6×10^{-4}	0.07
17	65	129	1.3×10^{-15}	6.1×10^{-16}	0.6
33	129	257	4.4×10^{-16}	—	3

Table 4.22: Results of the method of [13] for Example 3

n_{ka}	energy balance error	max. abs. error	time (sec)
400	2.6×10^{-1}	3.8×10^{-1}	0.004

Table 4.23: Result of the KA-based method for Example 3

n	e_n	this work's error	[13] method's error	KA-based method's error
-1	3.9×10^{-1}	-4.3×10^{-5}	-2.6×10^{-5}	1.2×10^{-1}
0	6.1×10^{-1}	1.3×10^{-4}	1.6×10^{-4}	-3.8×10^{-1}

Table 4.24: Efficiencies and errors for Example 3. The errors listed for this work as well as for the method of [13] come from the two solvers' "mod. acc." solutions.

shape of the scattering surface relative to the wavenumber of the incident field.

We compute the scattering efficiencies for these configurations using our approach, the method from [13] and the KA-based algorithm. Using the method of this thesis and that of [13], we generate solutions at two levels of precision as we did in Section 4.2.2.1; reference solutions by the solver of [13] were also computed so as to compare the accuracy of individual scattering efficiencies. While our method is somewhat slower than the approach of [13] in computing the more accurate solution for the $\frac{k}{2\pi} = 10.5$, $\theta = 85^\circ$ case, it is faster—sometimes dramatically so—in generating a solution of lesser accuracy for this case as well as all of the solutions of the other cases. Also, the KA-based solver is useful in obtaining moderately accurate results for the $\theta = 30^\circ$ cases (which have no multiple reflections), while for the $\theta = 85^\circ$ cases it breaks down due to presence of multiple reflections. The approach of this thesis, on the other hand, suffers no such breakdowns; furthermore, it yields solutions with similar or better accuracy in short computational times.

$\theta = 30^\circ$ **cases:** We first consider scattering from the multi-scale surface by an plane wave with incidence angle $\theta = 30^\circ$; plots containing this surface along with the direction of propagation vector for the incident field are given in Figure 4.8. Recalling the phase function

$$\phi_2(x, x') = \sqrt{(x - x')^2 + [f(x) - f(x')]^2} - (\sin(\theta), -\cos(\theta)) \cdot \big(x - x', f(x) - f(x')\big), \quad (4.3)$$

it can be shown by the test described in Section 2.3.3.2 that these scattering configurations only have simple reflections, since $\frac{\partial \phi_2(x,x')}{\partial x'} \neq 0$ for any x, x' (see Figure 4.9 for an example of this). Thus, for our method we compute $\mu_2(x)$; we also include results using $\mu_1(x)$ (the representation of the density more appropriate for problems with multiple scattering) by way of comparison. The scattering efficiencies for the $\frac{k}{2\pi} = 10.5$ case are plotted in Figure 4.10, and those of the $\frac{k}{2\pi} = 1000.5$ problem are depicted in Figure 4.11.

Our solver computes the efficiencies for these cases either to a moderate level of accuracy or to machine precision in very short times (Table 4.25). For example, the code only takes 21 seconds to compute machine-level-accurate efficiencies for the $\frac{k}{2\pi} = 10.5$ problem, and it takes even less time—11 seconds—for the $\frac{k}{2\pi} = 1000.5$ problem. Even if we compute the solutions using $\mu_1(x)$, the computational times are still quite short. Note that for $\frac{k}{2\pi} = 10.5$ there is no significant difference in using $\mu_1(x)$ instead of $\mu_2(x)$ since for this case

$n_t = n_i = 220$.

We solve these same problems using the approach of [13]; Table 4.26 lists the results of applying this method to these cases, with many sets of data for the $\frac{k}{2\pi} = 1000.5$ problem being given in order to demonstrate the length of time required to compute the solution of that case to near-machine-level accuracy. Comparing Tables 4.25 and 4.26, we find that the method of this thesis is noticeably faster than the method of [13] in solving these cases. In particular, when calculating $\mu_2(x)$ for the $\frac{k}{2\pi} = 1000.5$ case it is nearly 4100 times faster in computing the machine-level-accurate efficiencies. Furthermore, even though our method requires the same number of degrees of freedom for its solutions which are based upon $\mu_1(x)$ as the approach of [13] does for its solutions, it still takes less computational time for each problem—especially for the configuration with $\frac{k}{2\pi} = 1000.5$.

The KA-based method computes the efficiencies to a fair degree of accuracy in a very short amount of time for these cases. The energy balance errors and maximum absolute errors are given in Table 4.27. To further describe the accuracy of the computations, relative error plots (plots of the solutions' computed efficiencies minus the reference efficiencies divided by the reference efficiencies) are given in Figures 4.10 and 4.11; these figures include plots of all of the relative errors, but they also include relative error plots—denoted as "relative error (filtered)"—for only those reference efficiencies of more significant size (greater than 1×10^{-4}), since the relative errors for much smaller efficiencies (many efficiencies are as small as 1×10^{-30}) can be very large (e.g., 1000 or even much larger) yet, depending upon the particular application, may or may not pertain to the overall accuracy of the method. These plots indicate that in both cases the KA-based method computes at least the most significant efficiencies to within a few percent of their correct values.

For these cases, the method of this thesis is able to generate solutions of similar accuracy in similar short amounts of time (Table 4.25, Figures 4.10 and 4.11) compared to the KA-based approach, and for certain applications such solutions may be sufficiently accurate. If desired, it also is able to calculate rather quickly much more accurate solutions, e.g., the previously discussed machine-level-accurate results.

Remark 4.2.7. *The relative errors for our method's $\mu_1(x)$-based solution of the $\frac{k}{2\pi} = 10.5$ case are nearly identical to those of its $\mu_2(x)$-based solution, and the relative errors for its $\mu_1(x)$-based solution of the $\frac{k}{2\pi} = 1000.5$ case are very similar to those of the $n_{fl} = 401$ solution by the solver of [13].*

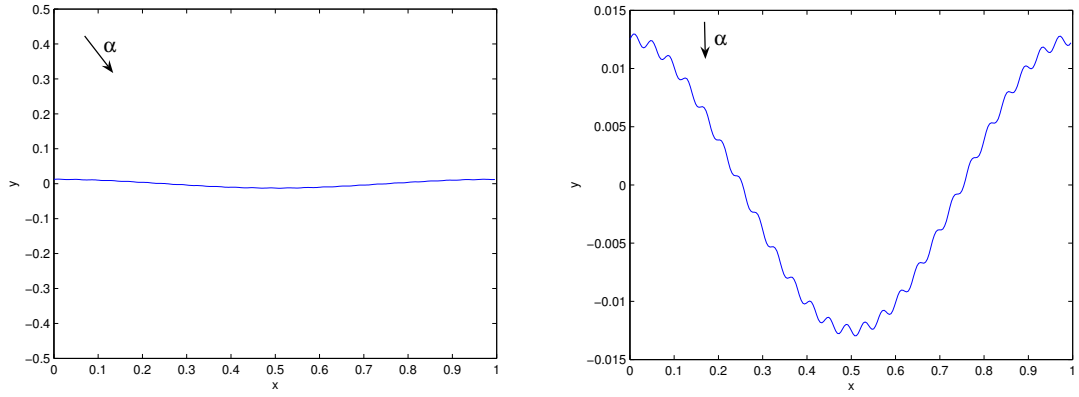


Figure 4.8: The multi-scale surface $y = f(x) = \frac{0.025}{2}\left[\cos(2\pi x) + 0.04\sin(50\pi x)\right]$ with incidence angle vector $\boldsymbol{\alpha} = (\sin(30^\circ), -\cos(30^\circ))$

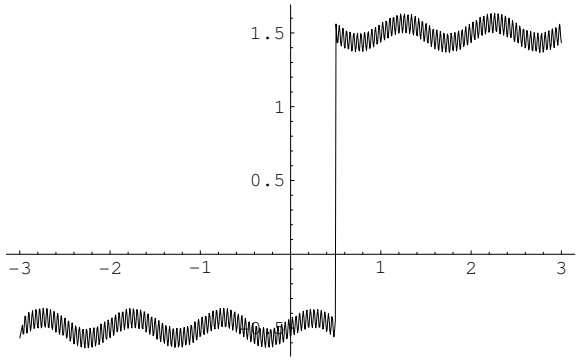


Figure 4.9: Plot of $\frac{\partial \phi_2(x,x')}{\partial x'}$ for $\theta = 30^\circ$ with $x = 0.5$ (the function is discontinuous at $x' = x$)

$\theta = 85^\circ$ **cases:** Scattering from the multi-scale grating by incident fields with $\theta = 85^\circ$ is also examined. Given the first derivative $f'(x) = \frac{0.025}{2}\left[-2\pi\sin(2\pi x) + 0.04 \times 50\pi\cos(50\pi x)\right]$ of the scattering profile (plotted in Figure 4.12), and since $\cot(85^\circ) \approx 0.0875$, shadowing is present according to the test of Section 2.3.3.3. Therefore, for our algorithm's computations the unknown $\mu_1(x)$ is used for the two cases under consideration. Figures 4.13 and 4.14 contain plots of the scattering efficiencies for the $\frac{k}{2\pi} = 10.5$ problem and the $\frac{k}{2\pi} = 1000.5$ problem, respectively. For these cases, the numerical parameters used for the various methods under consideration as well as the computational results achieved by these approaches are listed in Tables 4.28–4.30.

This thesis' approach still performs well even for these near-grazing configurations. Our method's total computational times are substantially higher for these problems than they

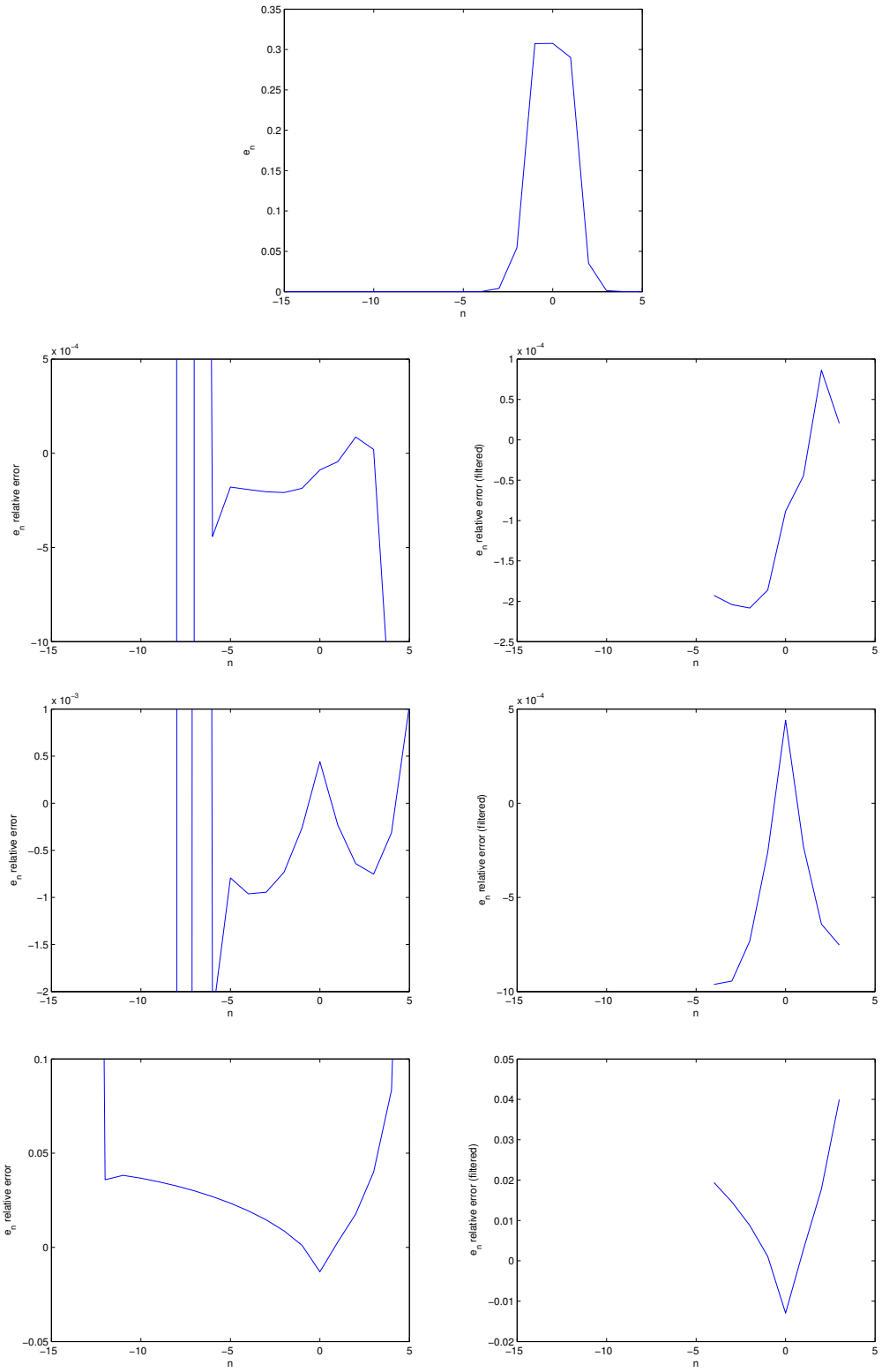


Figure 4.10: Efficiencies (top), relative errors of this work's "mod. acc." solution (second level), relative errors of the "mod. acc." solution produced by the method of [13] (third level) and relative errors of the KA-based method's solution (bottom) for the $\frac{k}{2\pi} = 10.5$, $\theta = 30^\circ$ multi-scale case

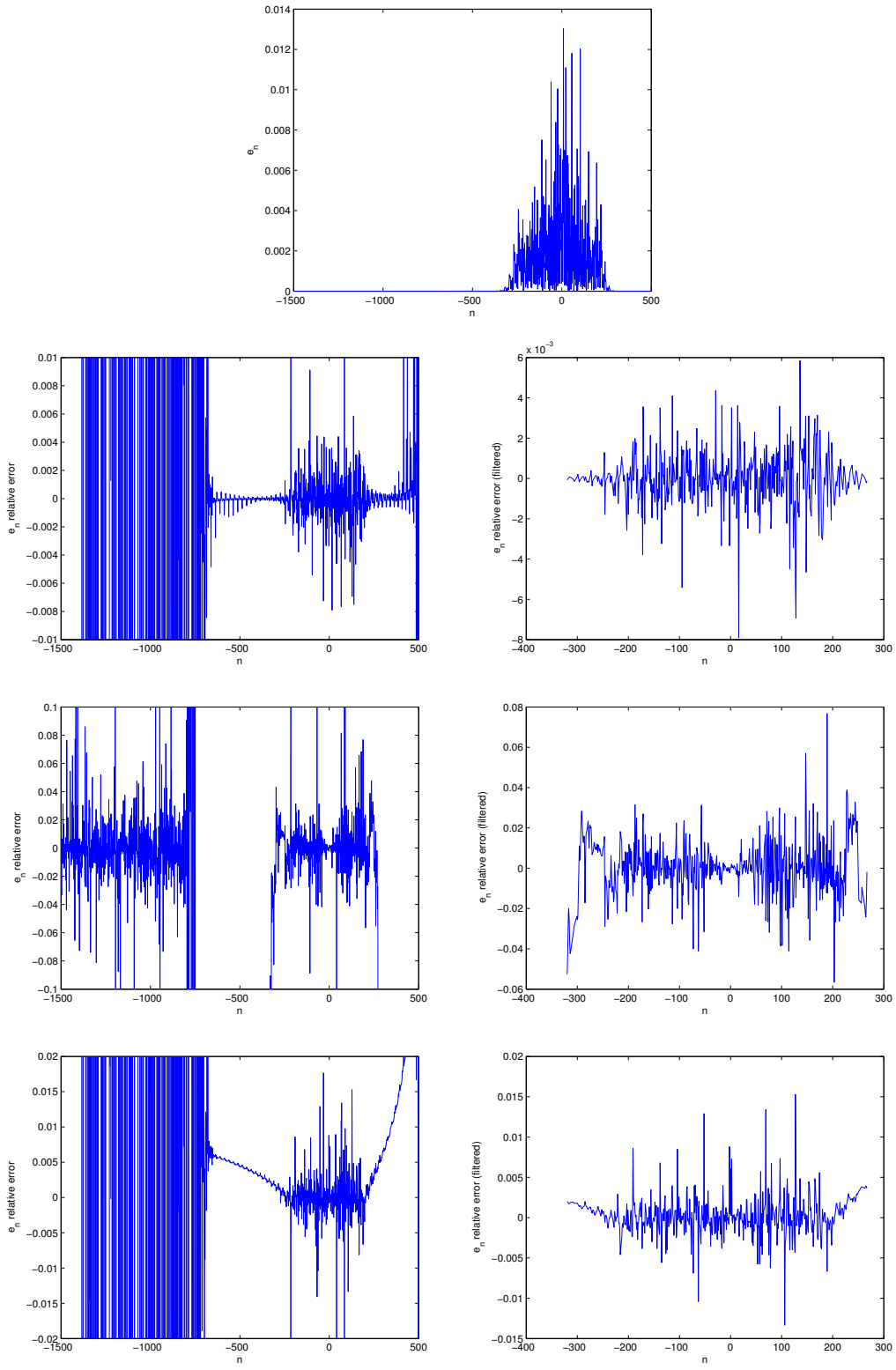


Figure 4.11: Efficiencies (top), relative errors of this work's "mod. acc." solution (second level), relative errors of the "mod. acc." solution produced by the method of [13] (third level) and relative errors of the KA-based method's solution (bottom) for the $\frac{k}{2\pi} = 1000.5,\ \theta = 30^\circ$ multi-scale case

rep.	$\frac{k}{2\pi}$	n_{t}	n_{i}	A_{sp}	A	e.b. error	max. abs. error	time (sec)
$\mu_2(x)$	10.5	64	64×1	0.875	2	1.1×10^{-4}	5.7×10^{-5}	0.07
$\mu_2(x)$	10.5	220	220×1	0.875	200	0.0×10^{-16}	1.3×10^{-15}	21
$\mu_2(x)$	1000.5	200	200×10	0.04375	0.05	1.3×10^{-6}	8.5×10^{-6}	1.5
$\mu_2(x)$	1000.5	600	600×8	0.4375	0.5	2.2×10^{-16}	1.4×10^{-15}	11
$\mu_1(x)$	10.5	64	64×1	0.875	2	1.1×10^{-4}	5.7×10^{-5}	0.07
$\mu_1(x)$	10.5	220	220×1	0.875	200	2.2×10^{-16}	1.3×10^{-15}	21
$\mu_1(x)$	1000.5	400	400×5	0.04375	0.05	4.3×10^{-6}	7.5×10^{-5}	1.6
$\mu_1(x)$	1000.5	1200	1200×4	0.4375	0.5	6.7×10^{-16}	1.9×10^{-15}	19

Table 4.25: Results for the multi-scale cases with $\theta = 30°$ using this work's method. The energy balance errors are listed as "e.b. error."

$\frac{k}{2\pi}$	n_{fl}	n_{ch}	n_{pg}	energy balance error	max. abs. error	time (sec)
10.5	65	65	33	7.9×10^{-5}	1.4×10^{-4}	2.9
10.5	221	257	129	8.2×10^{-14}	4.0×10^{-14}	48
10.5	441	513	257	2.2×10^{-16}	—	249
1000.5	401	4097	65	2.6×10^{-6}	6.7×10^{-5}	7401
1000.5	1201	4097	1025	8.0×10^{-8}	2.2×10^{-6}	22453
1000.5	1201	8193	33	1.5×10^{-5}	1.7×10^{-6}	44393
1000.5	1201	8193	1025	6.7×10^{-14}	—	44882

Table 4.26: Results of the method of [13] for the multi-scale cases with $\theta = 30°$

$\frac{k}{2\pi}$	n_{ka}	energy balance error	max. abs. error	time (sec)
10.5	2000	1.6×10^{-3}	4.0×10^{-3}	0.02
1000.5	2000	1.2×10^{-6}	1.5×10^{-5}	1.3

Table 4.27: Result of the KA-based method for the multi-scale cases with $\theta = 30°$

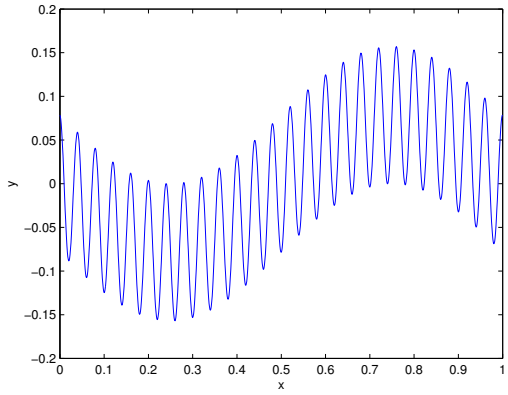


Figure 4.12: The multi-scale surface's first derivative

$\frac{k}{2\pi}$	n_t	n_i	A_{sp}	A	energy balance error	max. abs. error	time (sec)
10.5	64	64×1	0.875	200	1.2×10^{-4}	7.7×10^{-5}	1.8
10.5	220	220×1	0.875	6000	6.2×10^{-15}	2.4×10^{-14}	612
1000.5	200	200×12	0.875	30	1.0×10^{-4}	2.2×10^{-4}	34
1000.5	1200	1200×5	0.875	750	6.2×10^{-14}	1.6×10^{-13}	11992

Table 4.28: Results for the multi-scale cases with $\theta = 85°$ using this work's method ($\mu_1(x)$)

are for the $\theta = 30°$ problems, while the times for the solver of [13] are only moderately higher for $\frac{k}{2\pi} = 10.5$ and are basically the same for $\frac{k}{2\pi} = 1000.5$. In particular, for $\frac{k}{2\pi} = 10.5$ our solver is about 5 times slower than the method of [13] in computing machine-level-accurate efficiencies. Nevertheless, it is still faster than the solver of [13] in calculating such efficiencies for the $\frac{k}{2\pi} = 1000.5$ problem, and it is more efficient in computing "mod. acc." solutions for both cases. Additionally, the KA-based method computes the largest efficiency of the $\frac{k}{2\pi} = 10.5$ case to within about 1%, but for most of the efficiencies of both cases its results are rather inaccurate (Figures 4.13 and 4.14). Our method suffers no such breakdown, however, and its "mod. acc." solutions in particular—which have computational errors within about 1% for their significantly-sized efficiencies—are quickly computed.

4.2.2.3 Other Cases

Later in this thesis we will describe the results of applying our algorithm to a large number of additional cases that are systematically chosen in order to demonstrate the dependence of the numerical parameters of our solver upon key physical parameters (Section 4.3). All

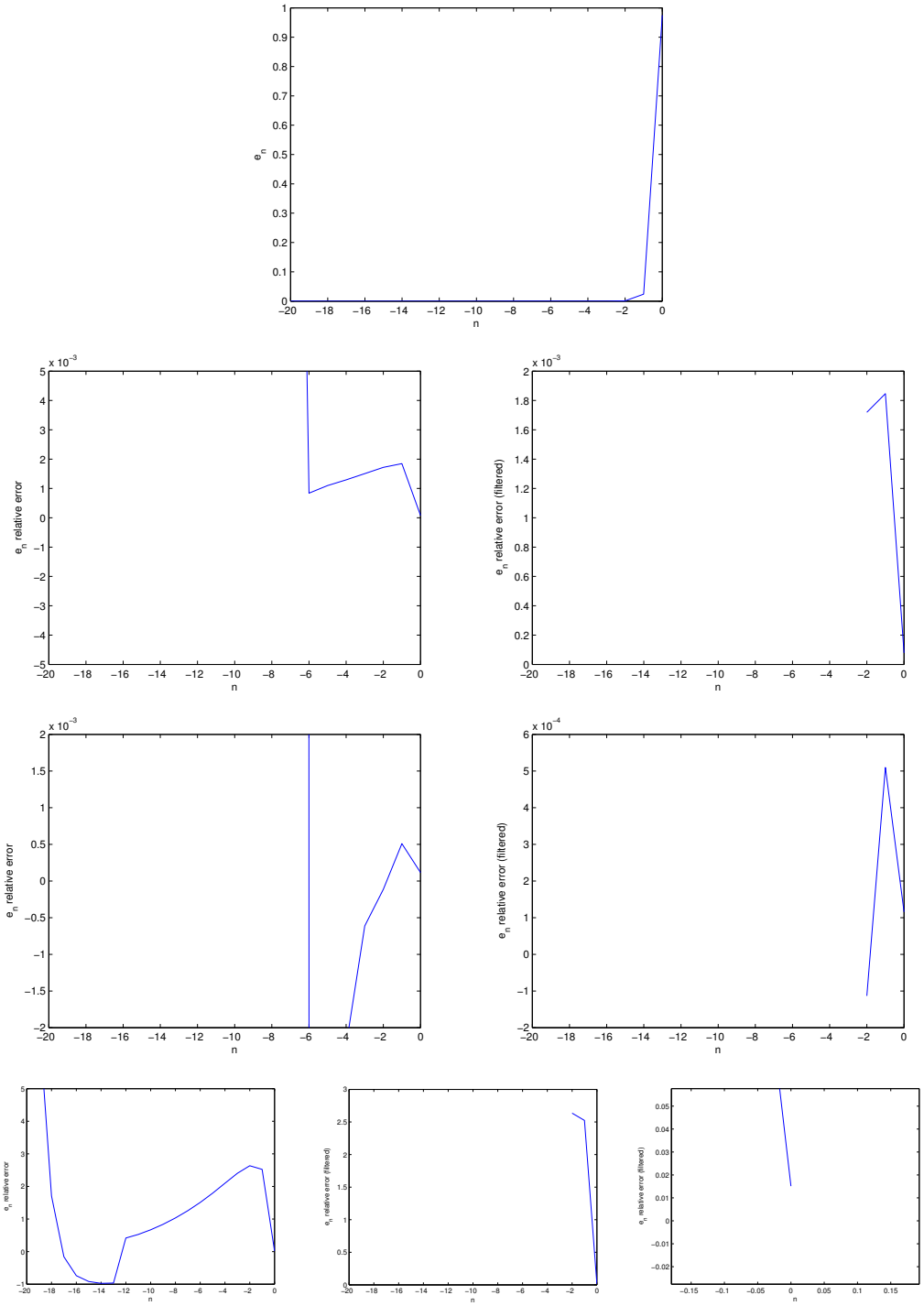


Figure 4.13: Efficiencies (top), relative errors of this work's "mod. acc." solution (second level), relative errors of the "mod. acc." solution produced by the method of [13] (third level) and relative errors of the KA-based method's solution (bottom) for the $\frac{k}{2\pi} = 10.5$, $\theta = 85°$ multi-scale case

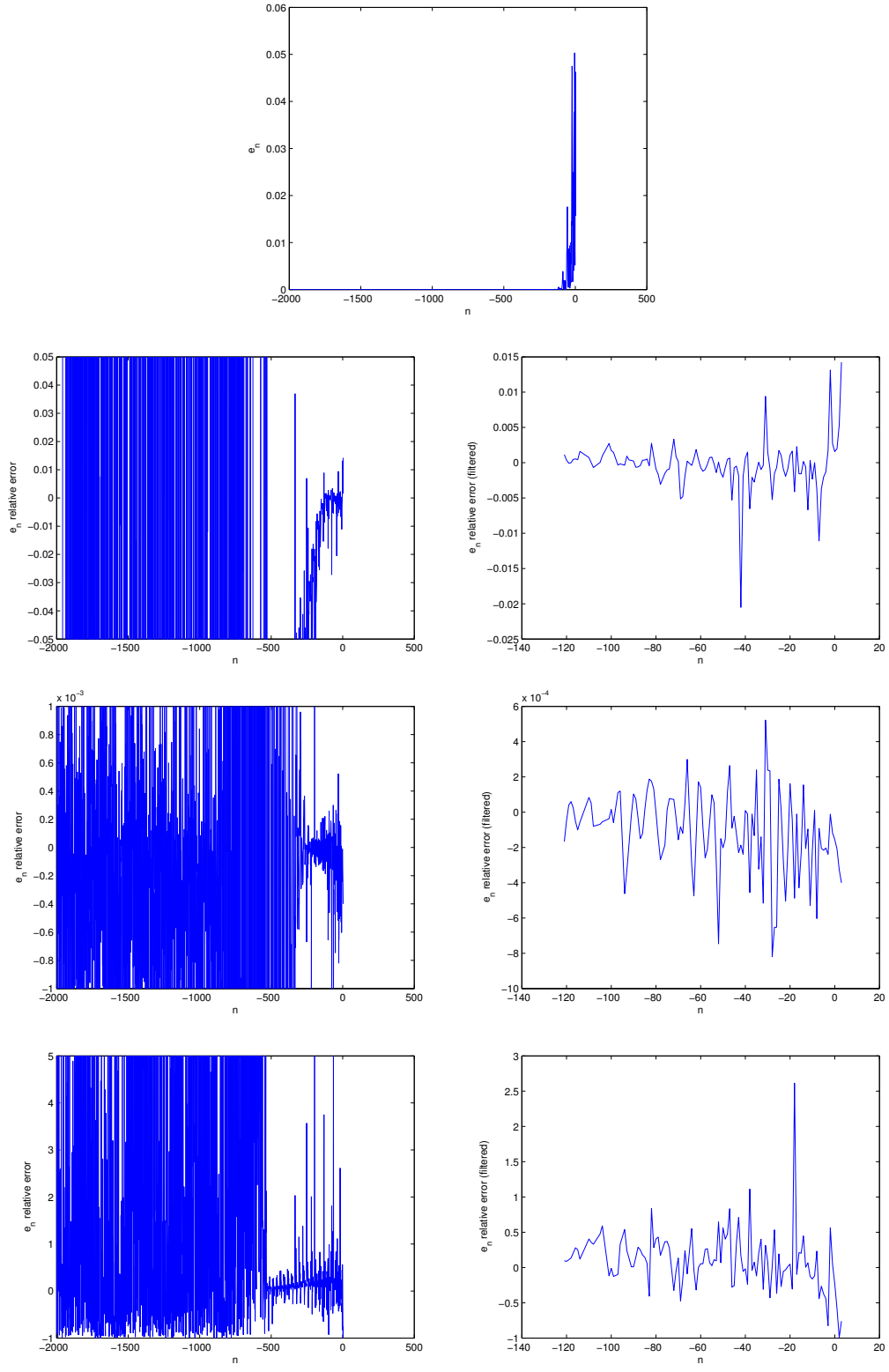


Figure 4.14: Efficiencies (top), relative errors of this work's "mod. acc." solution (second level), relative errors of the "mod. acc." solution produced by the method of [13] (third level) and relative errors of the KA-based method's solution (bottom) for the $\frac{k}{2\pi} = 1000.5,\ \theta = 85^\circ$ multi-scale case

$\frac{k}{2\pi}$	n_{fl}	n_{ch}	n_{pg}	energy balance error	max. abs. error	time (sec)
10.5	65	129	33	1.2×10^{-4}	1.1×10^{-4}	5.6
10.5	221	513	257	1.1×10^{-14}	6.7×10^{-16}	126
10.5	441	2049	513	1.1×10^{-14}	—	1467
1000.5	401	4097	129	1.6×10^{-4}	2.3×10^{-5}	7381
1000.5	1201	8193	1025	6.4×10^{-14}	—	44741

Table 4.29: Results of the method of [13] for the multi-scale cases with $\theta = 85°$

$\frac{k}{2\pi}$	n_{ka}	energy balance error	max. abs. error	time (sec)
10.5	2000	7.8×10^{-2}	6.0×10^{-2}	0.02
1000.5	4000	6.3×10^{-2}	1.9×10^{-2}	2.5

Table 4.30: Result of the KA-based method for the multi-scale cases with $\theta = 85°$

of those scattering problems involve surfaces of the form $f(x) = \frac{h}{2}\cos(2\pi x)$, and in that study we vary the height h, the wavenumber k and the incidence angle θ.

For completeness, we further compare the method of this thesis and the method of [13] by computing the efficiencies of some of those cases using both approaches. The results of all of the previous sections' examples which only contain simple reflections favor the method of this thesis in terms of total computational times. Thus, all of the systems chosen for this additional set of tests include the presence of multiple reflections; accordingly, the representation $\mu_1(x)$ is employed for all of these calculations by our algorithm. The KA-based approach was not applied to these problems, since (as shown in the previous sections) it does not perform well when multiple reflections are present.

The full details of the results for these cases are given in Section 4.3, but we here note the physical parameters of the cases of comparison as well as the computational times taken to compute the efficiencies of these problems at or near machine-level accuracy (Table 4.31). A clear pattern emerges: the approach of [13] is faster for the lower wavenumber (i.e., $\frac{k}{2\pi} = 10$) cases, while the method of this thesis is faster for the higher wavenumber (i.e., $\frac{k}{2\pi} = 100,\ 1000$) cases. Just as in the previous sections, our ability to reduce the integration window size A when applying this thesis' approach to increasingly large wavenumber problems is a key factor in keeping computational times relatively short.

h	$\frac{k}{2\pi}$	θ	this work's time (sec)	[13] method's time (sec)
0.25	100	10°	178	190
0.25	1000	10°	8030	62068
4.0	10	10°	467	37
4.0	10	60°	545	37
0.025	10	87°	82	3.9
0.025	1000	89°	456	2078

Table 4.31: Further computational results

4.2.3 Simulated 1-D Ocean Surfaces

In addition to considering deterministic surfaces, we also investigate scattering from a randomly generated periodic surface—a surface which corresponds to the surface waves of the deep ocean along one dimension. We compare the results of our approach to those resulting from the Kirchhoff approximation-based method, since the Kirchhoff approximation is useful for certain ocean scattering problems (see, e.g., [36] for a comparison of the method presented there with a KA-based method as they are applied to simulated ocean surfaces). We find that our method not only performs well for scattering at $\theta = 5°$ (where the KA-based approach also does well), but it also yields rather accurate results even at $\theta = 80°$ and $\theta = 85°$ (where the KA-based method breaks down).

To generate this random surface, we use a code provided by the Jet Propulsion Laboratory that is based upon the discussion found in [23]. In this paper, a directional wave spectrum model is developed for wind-driven surface waves of the ocean. The model equation—see [23, equation 67]—is dependent upon a few environmental parameters in addition to the two-dimensional wave vectors of the ocean waves. For the example considered here, we set the wind and dominant wave directions to be aligned ($\bar{\theta} = 0$), the sea to be "fully-developed" ($\Omega_c = 0.84$) and the wind speed at 10 m above the surface to be a moderate value ($U_{10} = 7\ m/s$); the "friction velocity" u^* is computed as a function of U_{10} in the code. We choose the direction for our one-dimensional surface realization to be along-wind ($\varphi = 0$). Given these values plus other parameters built into the model of [23], a one-dimensional power spectrum is computed for a 512 m-long periodic surface discretized at 2 m increments (i.e., a 256 point discretization), and Fourier coefficients for the surface are then randomly generated from this spectrum. See Figure 4.15 for the surface and its Fourier coefficients (the FFTs are scaled by the number of discretization points); we note that the

surface, while seemingly flat when plotted using a $1:1$ aspect ratio, contains small scale features which strongly affect the scattering of high-frequency waves (as we demonstrate below).

This surface is then prepared for use in our scattering code. First, it is interpolated using FFTs so as to be discretized at a much higher resolution. Then, the first and second derivatives of the scattering surface—required for our numerical quadratures (see Section 2.5)—are computed in Fourier space; see Figure 4.15 for these derivatives. When running our code, we read these data from files and periodically extend the surface for a sufficient number of periods $(2n_{\text{per}}+1)$ given the integration window size A.

For the incident field, we choose a wavelength λ that is similar to the wavelengths of GPS signals. GPS satellites transmit signals for civilian use at frequencies 1575.42 MHz (L1), 1227.60 MHz (L2) and, beginning in the year 2007, 1176.45 MHz (L5) [24]. Given the speed of light $c = 3\times10^8\ m/s$, the L2 frequency corresponds to $\lambda = \frac{3\times10^8\ m/s}{1.2276\times10^9\ 1/s} \approx \frac{1}{4}\ m$, so we use $\lambda = \frac{1}{4}$ for our wavelength (thus, the scattering surface has a 2048λ-length period and a somewhat larger arc length per period). Also, three incidence angles are considered: $\theta = 5^\circ$ (i.e., nearly normal incidence) along with $\theta = 80^\circ$ and 85° (i.e., grazing angles of 10° and 5°).

We use our $\mu_1(x)$-based method since scattering problems of this type can involve multiple scattering and shadowing, and in doing so a variety of numbers of target points per period n_{t} and integration window sizes A (modifying A_{sp} as needed) are taken. The number of integration points per period n_{i} is set to be the number of points per period in the discretization of the surface; for most cases, $n_{\text{i}} = 11520$, but for certain reference solutions of the $\theta = 80^\circ$ and 85° cases we use the larger value $n_{\text{i}} = 15360$. The KA-based approach is employed for various n_{ka}-point discretizations of its analytical approximation for $\mu_1(x)$.

The results of these computations are presented in Tables 4.32–4.37 and Figures 4.16–4.19. The tables indicate the energy balance errors achieved with the solutions. They also indicate the times (in seconds) required to compute the solutions; times taken without computing the efficiencies are listed as well, since often the computation of the 4096 efficiencies dominates much of the total time. To further describe the accuracy of the computations, the individual efficiencies are plotted using the reference solutions computed by our method, and, using these reference efficiencies, relative error and "relative error (filtered)" (Section 4.2.2.2) plots are given.

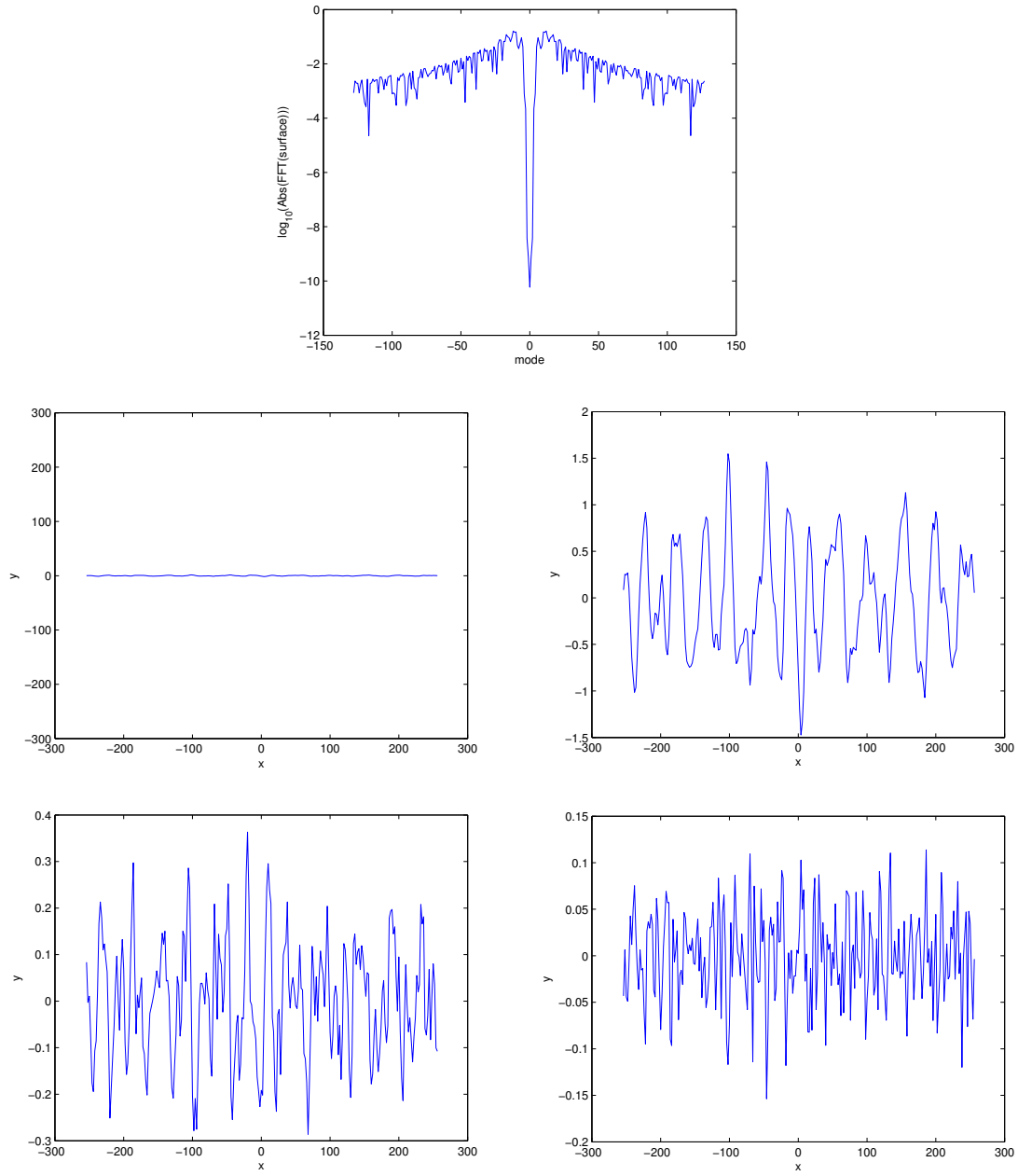


Figure 4.15: The magnitudes of the randomly generated Fourier coefficients (top), the simulated 1-D ocean surface (middle), the first derivative of the surface (bottom left) and the second derivative of the surface (bottom right)

For $\theta = 5°$, both our method and the KA-based method perform well, yielding accurate solutions in short computing times. Our method demonstrates convergence to machine precision in the energy balance error; the KA-based approach also has good values for the energy balance error, but, due to the non-convergent nature of the approximation, the values do not continue to decrease as the discretization increases (Tables 4.32 and 4.33). The KA-based computations are extremely quick, as expected, and the times for our approach are also quite small, especially given the levels of accuracy achieved. Additionally, using our method's $n_{\mathrm{t}} = 3840$ computation as a reference solution, individual efficiencies are plotted in Figure 4.16, and differences and relative errors are also plotted for the other two computations of our method as well as for the most accurate KA-based computation. The plot of the efficiencies indicates that the small scale features of the surface generate significant scattering in many directions besides the specular ($n = 0$) direction. The difference plots show that the accuracies indicated by the energy balance errors indeed hold on an efficiency-by-efficiency basis. Some of the relative errors of our method's $n_{\mathrm{t}} = 1280$ solution are somewhat large, but those for the more significantly sized efficiencies of the $n_{\mathrm{t}} = 2304$ solution are very small—corresponding to less than about 0.01% error for the more significantly sized efficiencies. For the KA-based approach, the relative errors are not quite as small, but they are still good (less than about 1% error for the larger efficiencies). For many applications, these errors of the KA-based solver may be sufficiently small. But, if necessary, our method is able to yield more accurate results than the KA-based method can, and it does so in short computational times.

Remark 4.2.8. *The efficiency-by-efficiency differences between our method's $n_t = 1280$ solution and the KA-based solver's $n_{ka} = 1280$ solution are much smaller than their errors relative to the reference solution; see Figure 4.17. Thus, the KA-based $n_{ka} = 1280$ solution's error and relative error plots are very similar to the plots for our method's $n_t = 1280$ solution and are not given here.*

For $\theta = 80°$ and $\theta = 85°$, however, the method of this thesis still performs well while the KA-based method is no longer accurate. Our approach yields energy balance errors of size 10^{-4} in 90 seconds and size 10^{-8} in a little over 90 minutes, but the best that the KA-based solver can do is compute solutions with energy balance errors of sizes 10^{-2} (for $\theta = 80°$) and 10^{-1} (for $\theta = 85°$); see Tables 4.34–4.37. Furthermore, taking our method's

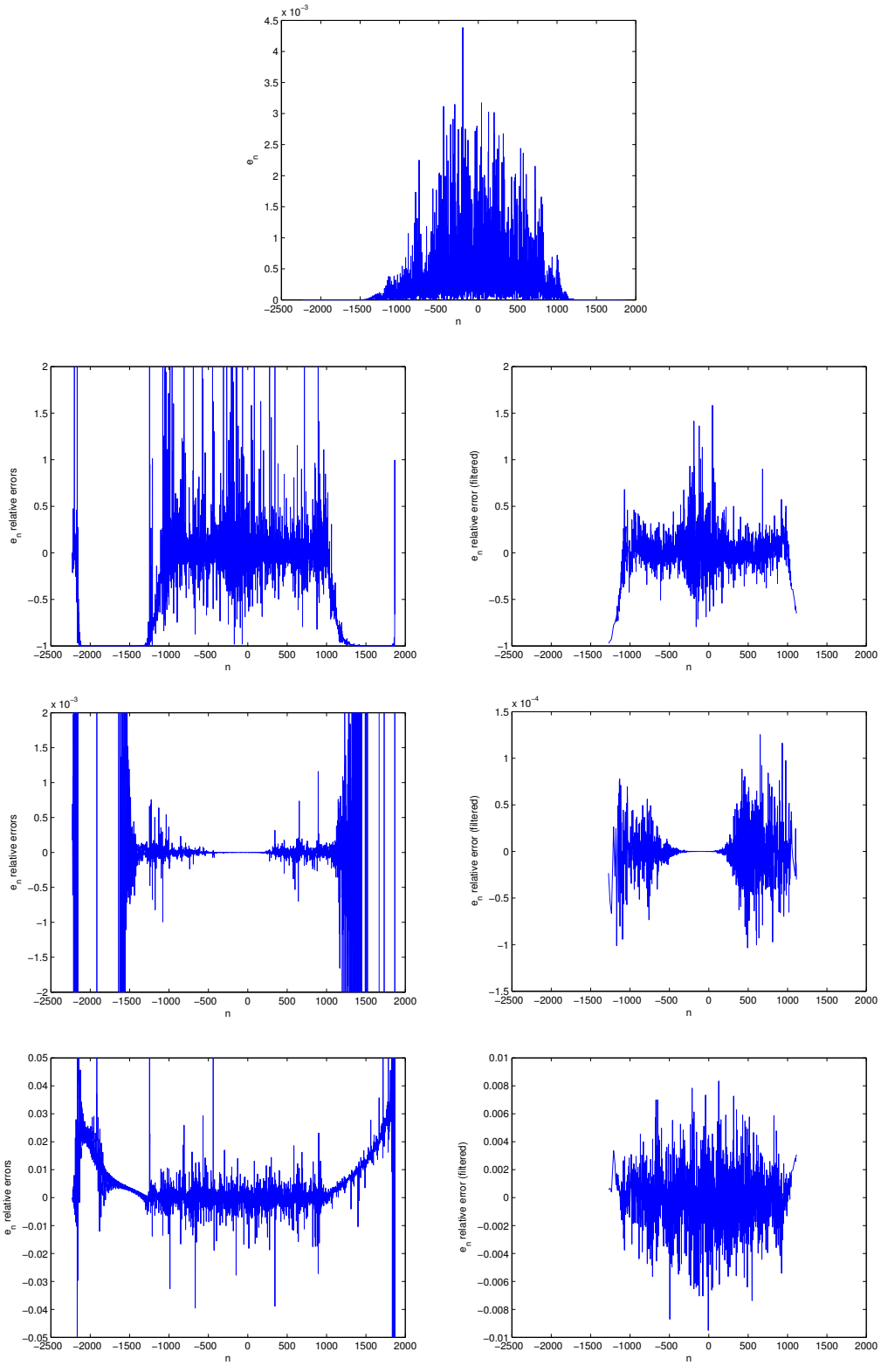


Figure 4.16: Efficiencies for the $\theta = 5°$ case (top), plus relative errors of our method's $n_{\mathrm{t}} = 1280$ solution (second level), our method's $n_{\mathrm{t}} = 2304$ solution (third level) and the KA-based approach's $n_{ka} = 3840$ solution (bottom)

n_{t}	n_{i}	A_{sp}	A	energy balance error	time (sec)	time w/o eff (sec)
1280	11520	0.000875	0.001	4.6×10^{-4}	17	2.2
2304	11520	0.0875	0.1	6.1×10^{-11}	32	17
3840	11520	0.0875	0.1	4.2×10^{-15}	43	29

Table 4.32: Table for the solutions for the $\theta = 5^{\circ}$ case as computed by the method of this thesis

n_{ka}	energy balance error	time (sec)	time w/o eff (sec)
1280	5.0×10^{-4}	15	0.06
2304	1.6×10^{-6}	15	0.06
3840	1.6×10^{-6}	16	0.07

Table 4.33: Table for the solutions for the $\theta = 5^{\circ}$ case as computed by the KA-based approach

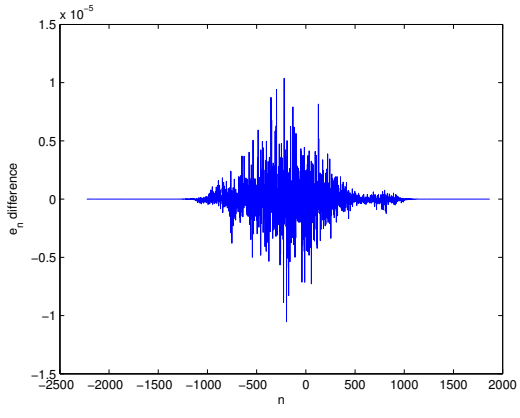


Figure 4.17: Efficiency-by-efficiency differences between our method's $n_{\mathrm{t}} = 1280$ solution and the KA-based $n_{ka} = 1280$ solution for the $\theta = 5^{\circ}$ case

n_t	n_i	A_{sp}	A	energy balance error	time (sec)	time w/o eff (sec)
768	11520	0.875	2	1.3×10^{-4}	90	75
1024	15360	0.875	150	4.4×10^{-8}	5555	—

Table 4.34: Table for the solutions for the $\theta = 80°$ case as computed by the method of this thesis. The time without computing the efficiencies was not determined for the $n_t = 1024$ reference case.

n_{ka}	energy balance error	time (sec)	time w/o eff (sec)
3840	4.3×10^{-2}	16	0.06

Table 4.35: Table for the solution for the $\theta = 80°$ case as computed by the KA-based approach

$n_t = 1024$ solutions (the 90 minute computations) as references, we see in Figures 4.18 and 4.19 that its $n_t = 768$ solutions (the 90 second computations) have most of their significant efficiencies correct to within 5% (a few efficiencies are about 10% off); the KA-based efficiencies, on the other hand, are wildly off. We note that for these low grazing angle cases the shadowing criterion of $f'(x) = -\cot(\theta)$ (see Section 2.3.3.3) is satisfied for certain x for $\theta = 80°$ $(-\cot(80°) \approx -0.18)$ and $\theta = 85°$ $(-\cot(85°) \approx -0.09)$, as can be seen in the first derivative plot in Figure 4.15, so there are both multiple reflections and shadowing occurring (see Section 2.3.3.1). Thus, the Kirchhoff approximation—which ignores such phenomena in its approximation—fails in these cases. The computational approach of this thesis, however, not only suffers no such difficulties, but it actually computes reasonably accurate results in only 90 seconds.

n_t	n_i	A_{sp}	A	energy balance error	time (sec)	time w/o eff (sec)
768	11520	0.875	2	6.6×10^{-4}	90	75
1024	15360	0.875	150	6.5×10^{-8}	5556	—

Table 4.36: Table for the solutions for the $\theta = 85°$ case as computed by the method of this thesis. The time without computing the efficiencies was not determined for the $n_t = 1024$ reference case.

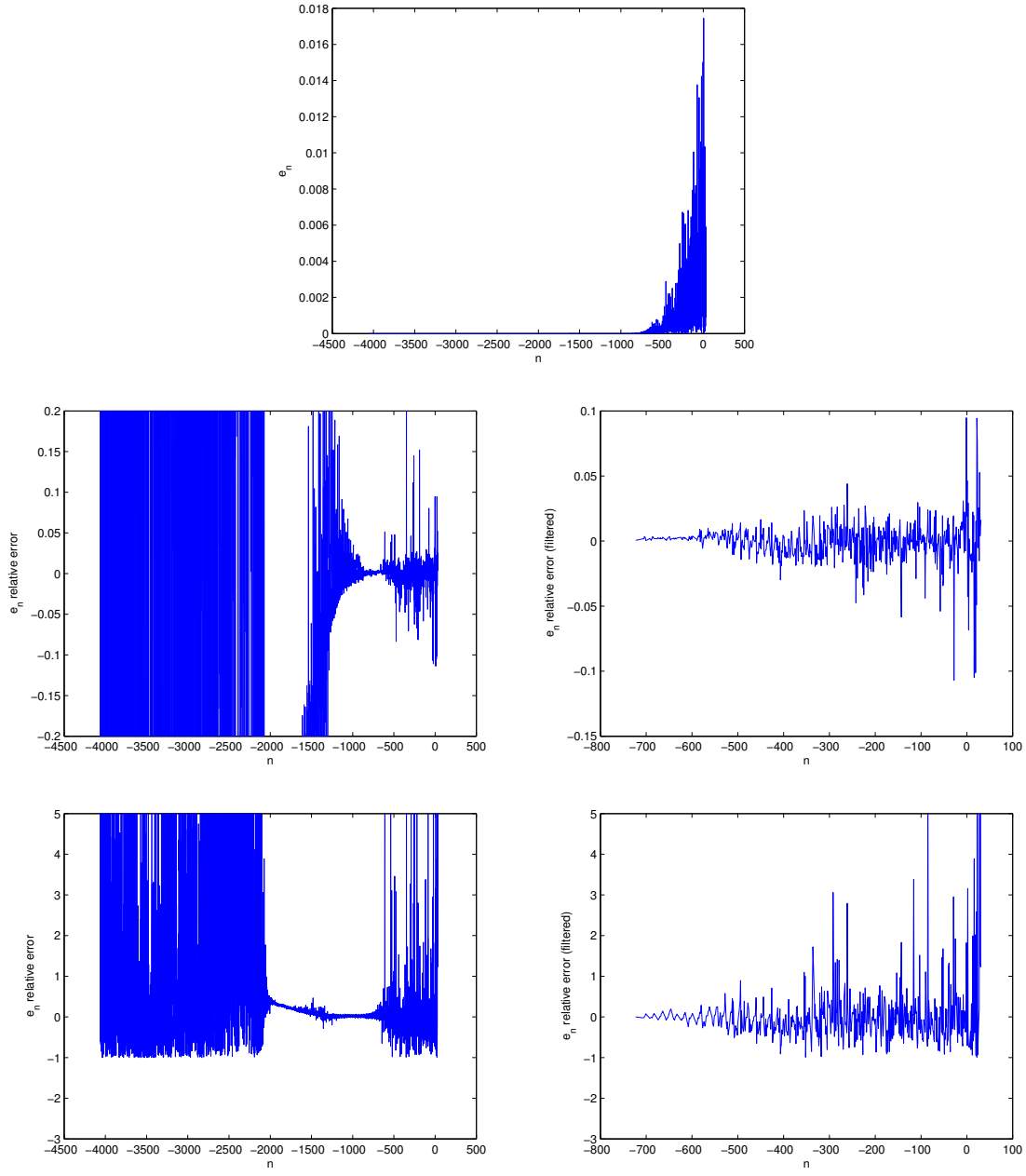


Figure 4.18: Efficiencies for the $\theta = 80°$ case (top), plus relative errors of our method's $n_{\mathrm{t}} = 768$ solution (middle) and the KA-based approach's $n_{ka} = 3840$ solution (bottom)

n_{ka}	energy balance error	time (sec)	time w/o eff (sec)
3840	1.6×10^{-1}	16	0.07

Table 4.37: Table for the solution for the $\theta = 85°$ case as computed by the KA-based approach

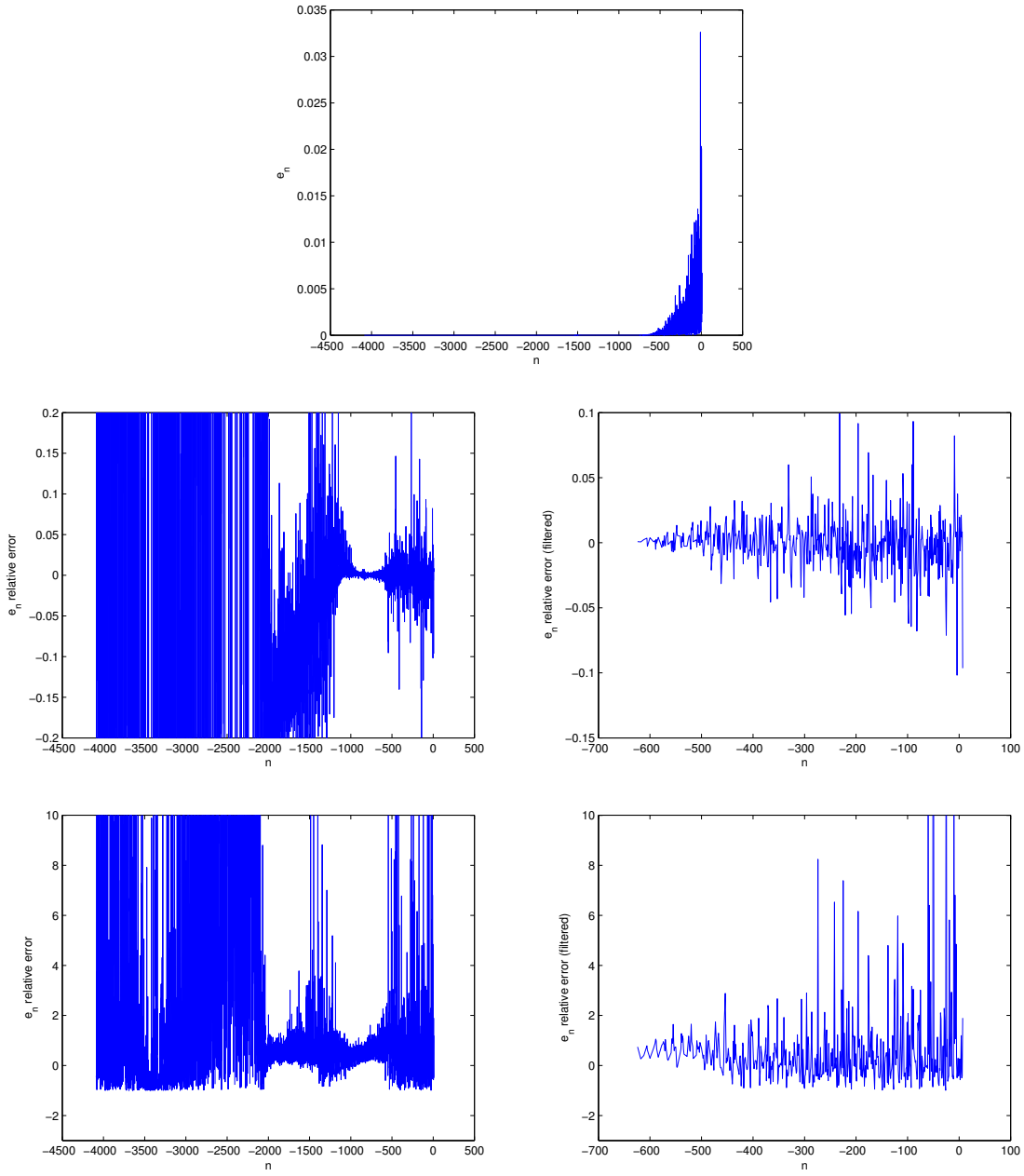


Figure 4.19: Efficiencies for the $\theta = 85°$ case (top), plus relative errors of our method's $n_{\mathrm{t}} = 768$ solution (middle) and the KA-based approach's $n_{ka} = 3840$ solution (bottom)

4.3 Dependence of Algorithm Parameters upon Physical Parameters

Using a grating profile of the form $f(x) = \frac{h}{2}\cos(2\pi x)$, we vary the height h and the incident field's wavenumber k and angle θ in order to investigate the dependence of the method of this thesis upon these physical quantities. By systematically examining a broad range of cases in this manner, data are generated which help serve to guide parameter choices even for scattering cases we do not consider in this thesis. Additionally, these data could be used in future work for purposes of case-by-case comparison with modified versions of our computational algorithm as well as with other numerical methods. Finally, certain key properties of our solver are illustrated and confirmed by the examples of this section.

4.3.1 Dependence upon k

We first vary the wavenumber k for certain values of h, fixing $\theta = 10^\circ$. The wavenumbers are chosen so that $\frac{k}{2\pi} = 10^n$ for various integers n. As discussed in Section 2.3.3.4, for gratings of the form considered in this section we have explicitly determined regions in (h,θ) space corresponding to configurations for which only simple reflections occur, for which there are also multiple reflections but not shadowing, and for which there are multiple reflections and shadowing; the boundaries of these regions are displayed in Figure 2.10. Here, two values of h are chosen for which only simple reflections occur given $\theta = 10^\circ$: one value far below the threshold and another value near (but below) it. Also, one value of h just above the threshold—for which there are also multiple reflections—is considered. By selecting h in this way, we show that there is a clear difference between the multiple-reflection cases and the simple-reflection cases in terms of the dependence of the numerical parameters upon k.

Remark 4.3.1. *Most of the values of k chosen for these problems are well away from Wood Anomaly values (Remark 2.1.5). A discussion of cases where k is at or near certain Wood Anomaly values is given in Section 4.3.4.*

For each of the scattering cases, we determine the minimum number n_t of target points per period, the minimum number n_i of integration points per period and the minimum integration window size A that are necessary to represent numerically all of the significant Fourier modes of the solution (either $\mu_1(x)$ or $\mu_2(x)$) as well as to achieve energy balance

errors at or near machine precision levels. Two different patterns in the values of these parameters emerge in this study. For the two sets of simple-reflection cases, there are values of k in the "resonance" regime for which the values of n_t (computing $\mu_2(x)$, as is appropriate for this regime) are at their maximum; the values of n_t are smaller for smaller and larger k. Additionally, as k continues to increase for these two sets of cases, n_i becomes directly proportional to k while A becomes inversely proportional to k. Thus, n_t and $n_i \times A$ remain fixed for these increasing k and the times to compute the solutions $\mu_2(x)$ are virtually constant. For the multiple-reflection cases, however, both n_t and n_i (computing $\mu_1(x)$, as is appropriate for this regime) increase without bound as do the times to compute $\mu_1(x)$ (even though A decreases as k increases). In addition to these results, the advantages of using the appropriate representations (either $\mu_1(x)$ or $\mu_2(x)$) for the solutions of the two sets of cases for which h is near the multiple reflections threshold are demonstrated.

4.3.1.1 Simple-Reflection Cases

$h = 0.025$ cases: The first scattering profile we consider has height $h = 0.025$. Setting the incidence angle to be $\theta = 10°$, we examine configurations containing only simple reflections; this choice of (h,θ) is well below the multiple reflection threshold in the plot found in Figure 2.10. The wavenumber k is varied so that $\frac{k}{2\pi} = 10^n$ for $n = -2, \ldots, 4$, and for each case the computational parameters n_t, n_i and A are determined as previously described (A_{sp} is reduced for those cases where $A < 1$). Here $\mu_2(x)$ is computed for each problem since there are no multiple reflections.

The data for these cases are given in Table 4.38, including the numerical parameters employed, the energy balance errors and the times needed to compute the solutions $\mu_2(x)$ without subsequently computing the scattering efficiencies. The values of n_t are 16 for low and high wavenumbers, but n_t increases to 38 for $\frac{k}{2\pi} = 10$ (in the "resonance" wavenumber regime). The integration window size A decreases as k increases, which is consistent with both our physical intuition (nearby scattering interactions dominate in the "high-frequency" wavenumber regime) as well as the case study of Section 2.2.3. In particular, for the "high-frequency" cases n_i is directly proportional to k while n_t and $n_i \times A$ are fixed (as discussed in Section 3.2.2), so that the times to compute the solutions $\mu_2(x)$ are essentially constant.

Remark 4.3.2. *The total computational times—including computing all of the scattering efficiencies using the usual trapezoidal quadrature rule—grow quadratically in k for the*

$\frac{k}{2\pi}$	n_{t}	n_{i}	A_{sp}	A	energy balance error	time w/o eff (sec)
$\frac{1}{100}$	16	16×3	0.875	8000	1.1×10^{-16}	14
$\frac{1}{10}$	16	16×4	0.875	1400	6.7×10^{-16}	3.2
1	24	24×4	0.875	900	4.4×10^{-16}	4.8
10	38	38×1	0.875	30	4.4×10^{-16}	0.13
100	16	16×20	0.875	2	4.4×10^{-16}	0.14
1000	16	16×200	0.175	0.2	1.8×10^{-15}	0.24
10000	16	16×2000	0.0175	0.02	7.8×10^{-16}	0.38

Table 4.38: Table for $\theta = 10^\circ$ and $h = 0.025$

larger wavenumber problems, since n_i grows linearly and since the number of scattering efficiencies to be computed grows linearly. For example, our solver takes 200 *seconds in total computational time for the case with $\frac{k}{2\pi} = 10000$. In practice, not all of the scattering efficiencies for such a case would necessarily be computed; instead, depending upon the physical phenomena being modeled, perhaps only values of the field $\psi^{scat}(\boldsymbol{r})$ at certain points near the scattering surface or only certain efficiencies would be required.*

We note that for $\frac{k}{2\pi} = \frac{1}{100}, \frac{1}{10}$ there is only the scattering efficiency $e_0 = 1$ for each case, making it unnecessary to solve for $\mu_2(x)$ first for these problems if accurate computations of the scattering efficiencies are the ultimately desired result. However, accurately computing $\mu_2(x)$ in this "low-frequency" regime may be desired for certain applications, and doing this requires increasing computational costs for our method as k decreases.

$h = 0.2$ cases: We next let $h = 0.2$ while keeping $\theta = 10^\circ$, and we vary k as before. These scattering configurations still only give rise to simple reflections, but now (h,θ) is near the multiple reflection threshold in Figure 2.10.

Qualitatively, the computational parameters for the solutions of these cases follow a similar pattern as those for the solutions of the $h = 0.025$ cases. See Table 4.39 for the results. Here the maximum value for n_{t} is 400 (for $\frac{k}{2\pi} = 100$); this is larger than $n_{\mathrm{t}} = 38$ (for $\frac{k}{2\pi} = 10$), the maximum value for the $h = 0.025$ cases, which is to be expected since the grating is deeper here. Again, $\mathcal{O}(1)$ computational time to evaluate the solutions $\mu_2(x)$ for the "high-frequency" cases is demonstrated.

The Fourier amplitudes of the solutions $\mu_2(x)$ of the $\frac{k}{2\pi} = 100$ and $\frac{k}{2\pi} = 1000$ configurations are plotted in Figure 4.20. These plots demonstrate the "resonance" at $\frac{k}{2\pi} = 100$: it is necessary to use $n_{\mathrm{t}} = 400$ target points per period for the $\frac{k}{2\pi} = 100$ case while $n_{\mathrm{t}} = 24$ is

$\frac{k}{2\pi}$	n_t	n_i	A_{sp}	A	energy balance error	time w/o eff (sec)
$\frac{1}{100}$	30	30×4	0.875	21000	4.4×10^{-16}	163
$\frac{1}{10}$	30	30×4	0.875	2100	2.2×10^{-16}	16
1	32	32×4	0.875	1500	0.0×10^{-16}	14
10	82	82×2	0.875	500	7.5×10^{-15}	14
100	400	400×3	0.875	200	7.3×10^{-15}	203
1000	24	24×250	0.875	1	4.8×10^{-15}	2.2
10000	24	24×2500	0.0875	0.1	2.2×10^{-14}	2.5

Table 4.39: Table for $\theta = 10^\circ$ and $h = 0.2$

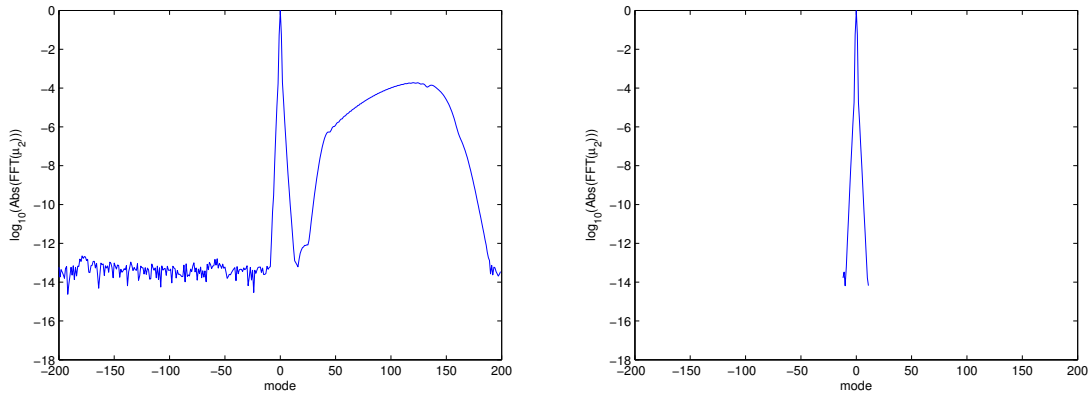


Figure 4.20: Fourier amplitudes for $\frac{k}{2\pi} = 100$ (left) and $\frac{k}{2\pi} = 1000$ (right) for $\theta = 10^\circ$ and $h = 0.2$

sufficient for the $\frac{k}{2\pi} = 1000$ case. Also, the substantial savings in using $\mu_2(x)$ (appropriate for these problems since they only contain simple reflections) instead of $\mu_1(x)$ in the computation of the $\frac{k}{2\pi} = 1000$ problem is shown in Table 4.40 and Figure 4.21; many fewer target points per period are needed for the solution $\mu_2(x)$ than are needed for $\mu_1(x)$, resulting in a much shorter total evaluation time for this case.

rep.	n_t	n_i	A_{sp}	A	e.b. error	max. abs. difference	time w/o eff (sec)
$\mu_2(x)$	24	24×250	0.875	1	4.8×10^{-15}	4.8×10^{-15}	2.2
$\mu_2(x)$	2000	2000×3	0.875	1	1.7×10^{-15}	4.1×10^{-15}	66
$\mu_1(x)$	2000	2000×3	0.875	1	8.9×10^{-16}	—	65

Table 4.40: Table for $\frac{k}{2\pi} = 1000$, $\theta = 10^\circ$ and $h = 0.2$. The energy balance errors are listed as "e.b. error."

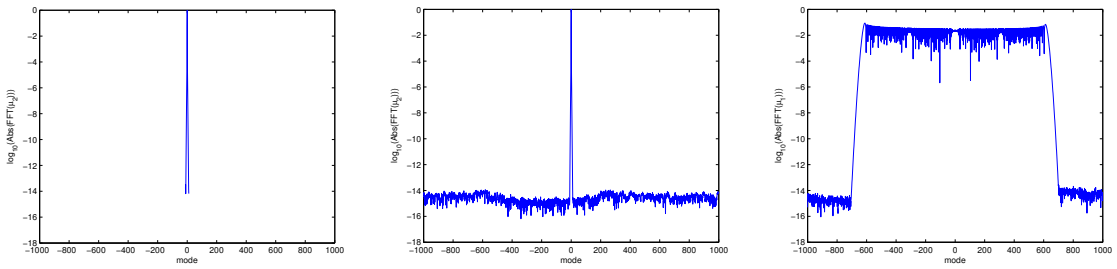


Figure 4.21: Fourier amplitudes for the $\mu_2(x)$, $n_t = 24$ representation (left), $\mu_2(x)$, $n_t = 2000$ representation (middle) and the $\mu_1(x)$ representation (right) for $\frac{k}{2\pi} = 1000$, $\theta = 10^\circ$ and $h = 0.2$

4.3.1.2 Multiple-Reflection Cases

In order to examine multiple-reflection cases, we set $h = 0.25$ and $\theta = 10^\circ$. Again, (h,θ) is near the multiple reflection threshold in Figure 2.10, although it now is just above it. For these problems we compute $\mu_1(x)$ in accordance with the presence of multiple reflections.

The results of these cases are significantly different from those of the simple-reflection cases. Even though A decreases and n_i increases as k increases, which also occurs for the simple-reflection cases, n_t increases without bound as k increases (Table 4.41), unlike for the previous cases where n_t remains constant in the "high-frequency" regime. Additionally, the computational time to evaluate the solution $\mu_1(x)$ continues to noticeably increase for "high-frequency" wavenumbers instead of leveling off. The total computational times (including computing the scattering efficiencies) for the $\frac{k}{2\pi} = 100$ and $\frac{k}{2\pi} = 1000$ cases are about the same as the times for computing only the solutions $\mu_1(x)$: 178 seconds and 8030 seconds, respectively.

Remark 4.3.3. *The total computational times required by the method of [13] to compute essentially machine-level accurate efficiencies are* 190 *seconds for the* $\frac{k}{2\pi} = 100$ *case and* 62068 *seconds for the* $\frac{k}{2\pi} = 1000$ *case (Table 4.43).*

As stated earlier, $\mu_1(x)$ is the appropriate solution for these cases given the presence of multiple reflections. Table 4.42 and Figure 4.22 demonstrate the approximately factor of 2 computational savings obtained by using $\mu_1(x)$ rather than $\mu_2(x)$ for the $\frac{k}{2\pi} = 1000$ problem.

$\frac{k}{2\pi}$	n_{t}	n_{i}	A_{sp}	A	energy balance error	time w/o eff (sec)
$\frac{1}{100}$	32	32×3	0.875	20000	3.3×10^{-16}	132
$\frac{1}{10}$	32	32×3	0.875	2000	4.4×10^{-16}	13
1	32	32×3	0.875	1400	1.8×10^{-15}	10
10	72	72×2	0.875	550	3.8×10^{-14}	12
100	340	340×2	0.875	350	2.1×10^{-14}	178
1000	3000	3000×2	0.875	200	1.8×10^{-13}	8030

Table 4.41: Table for $\theta = 10^\circ$ and $h = 0.25$

rep.	n_{t}	n_{i}	A_{sp}	A	energy balance error	time w/o eff (sec)
$\mu_1(x)$	3000	3000×2	0.875	200	1.8×10^{-13}	8030
$\mu_2(x)$	5800	5800×1	0.875	200	1.3×10^{-13}	15318

Table 4.42: Table for $\frac{k}{2\pi} = 1000$, $\theta = 10^\circ$ and $h = 0.25$

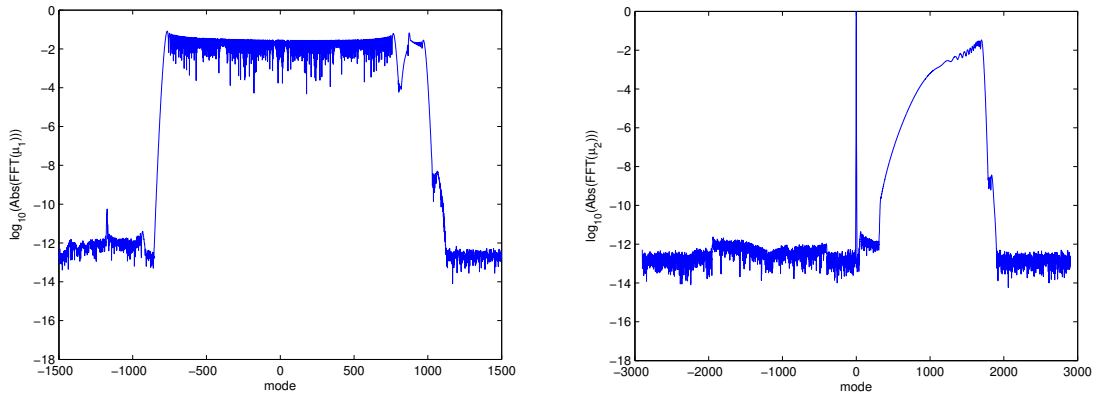


Figure 4.22: Fourier amplitudes for the $\mu_1(x)$ representation (left) and the $\mu_2(x)$ representation (right) for $\frac{k}{2\pi} = 1000$, $\theta = 10^\circ$ and $h = 0.25$

$\frac{k}{2\pi}$	n_{fl}	n_{ch}	n_{pg}	energy balance error	time (sec)
100	341	1025	257	5.1×10^{-15}	190
1000	3001	4097	1025	7.5×10^{-5}	30827
1000	3001	8193	1025	2.6×10^{-14}	62068

Table 4.43: Results of the method of [13] ($\theta = 10^\circ$ and $h = 0.25$)

4.3.2 Dependence upon h

Now, we vary the height h for $\theta = 10^\circ$ and $\theta = 60^\circ$, and we leave k fixed so that $\frac{k}{2\pi} = 10$ (k is in the "resonance" regime, and it is well away from being a Wood Anomaly value given the choices of θ). Re-using some of the $\theta = 10^\circ$ data from the wavenumber study of Section 4.3.1, we choose values of h from 0.025 to 4.0 and compute the efficiencies for the resulting scattering cases. The transition to the multiple reflections regime for $\theta = 10^\circ$ occurs at around $h = 0.225$, while for $\theta = 60^\circ$ it occurs at around $h = 0.075$ (Figure 2.10). Accordingly, we use $\mu_1(x)$ for the multiple-reflection cases and $\mu_2(x)$ for the simple-reflection cases. We also compute the simple-reflection cases using $\mu_1(x)$; additional physical insight is gained by generating data for all of the cases using just this representation for the unknown densities.

The number n_t of target points per period, the number n_i of integration points per period and the integration window size A tend to increase as h increases in order to maintain a high level of accuracy for both the $\theta = 10^\circ$ cases and $\theta = 60^\circ$ cases. This is true when using $\mu_1(x)$ and $\mu_2(x)$ in their typical regimes (Tables 4.44 and 4.46) as well as when using $\mu_1(x)$ for all of the cases (Tables 4.45 and 4.47). Accordingly, the computational times for these problems increase as h increases; our solver requires 467 seconds (about 8 minutes) to compute the efficiencies of the $h = 4.0,\ \theta = 10^\circ$ case to nearly machine precision level accuracy, and the $h = 4.0,\ \theta = 60^\circ$ case takes 545 seconds (about 9 minutes).

Remark 4.3.4. *As indicated in Table 4.48, the total computational time of the method of [13] for each of the $h = 4.0$ cases is only 37 seconds.*

Occasionally, however, n_i decreases as h increases. One reason for this is that n_i is an integer multiple of n_t, so there are some cases in which it is somewhat larger than the minimum required for machine precision accuracy. For example, $n_\mathrm{i} = 30 \times 2 = 60$ for the solution of the $h = 0.025,\ \theta = 10^\circ$ case using $\mu_1(x)$, while $n_\mathrm{i} = 40 \times 1 = 40$ for that of the $h = 0.05,\ \theta = 10^\circ$ case (Table 4.45). However, if we let $n_\mathrm{t} = 38$, $n_\mathrm{i} = 38 \times 1$ and $A = 30$, then the $\mu_1(x)$ solution of the $h = 0.025,\ \theta = 10^\circ$ case is machine-level accurate (with an energy balance error of 1.1×10^{-16})—as expected given the $\mu_2(x)$ solution described in Table 4.44.

Furthermore, for the sets of data which are based upon using either $\mu_1(x)$ or $\mu_2(x)$ in accordance the types of scattering that are occurring, n_t (and n_i) decreases for the $\theta = 10^\circ$

rep.	h	n_t	n_i	A	energy balance error	$\frac{n_t}{(s/\lambda)}$
$\mu_2(x)$	0.025	38	38×1	30	4.4×10^{-16}	3.8
$\mu_2(x)$	0.05	44	44×1	80	3.3×10^{-16}	4.4
$\mu_2(x)$	0.1	60	60×1	250	1.2×10^{-14}	5.9
$\mu_2(x)$	0.2	82	82×2	500	7.5×10^{-15}	7.5
$\mu_1(x)$	0.25	72	72×2	550	3.8×10^{-14}	6.3
$\mu_1(x)$	0.5	96	96×2	600	2.9×10^{-13}	6.6
$\mu_1(x)$	1.0	128	128×2	700	3.6×10^{-13}	5.6
$\mu_1(x)$	2.0	220	220×2	750	1.7×10^{-12}	5.3
$\mu_1(x)$	4.0	360	360×2	800	4.6×10^{-13}	4.4

Table 4.44: Table for various h ($\frac{k}{2\pi} = 10$, $\theta = 10°$)

cases as h increases from 0.2 to 0.25 (Table 4.44), and it (along with n_i) remains fixed as h increases from 0.05 to 0.1 for the $\theta = 60°$ problems (Table 4.46). In each of these instances the representation of the unknown density changes from $\mu_2(x)$ to $\mu_1(x)$. On the other hand, n_t always increases as h increases within the $\mu_1(x)$-only data (Tables 4.45 and 4.47), which indicates that the anomalies in the other data are due to the changes of representation.

Remark 4.3.5. *Some scattering configurations with only simple reflections have slightly lower n_t values for their solutions $\mu_1(x)$ than for the corresponding $\mu_2(x)$—the opposite of the usual pattern (see, e.g., Sections 2.3.3.5). These cases have a somewhat small wavenumber ($k = 10 \times 2\pi$); in Section 4.3.1 we showed that cases which have the same gratings (either with $h = 0.2$ or 0.25) and incidence angle $\theta = 10°$ but with $k = 1000 \times 2\pi$ follow the general rule of $\mu_2(x)$ requiring fewer degrees of freedom when there are no multiple reflections (for $h = 0.2$) and $\mu_1(x)$ requiring fewer degrees of freedom when multiple reflections do arise (for $h = 0.25$).*

Since for these cases the incident field is fixed (k and θ are fixed) while the scattering surface is varied, we also consider the number $\frac{n_t}{(s/\lambda)}$ of degrees of freedom per wavelength (Section 4.2.2.1; see in particular Remark 4.2.4) for each solution. This quantity is maximized within the range of heights h considered in this section, both for $\theta = 10°$ and for $\theta = 60°$. There is more than one local maximum in each of the data sets of Tables 4.44 and 4.46 (based upon either $\mu_1(x)$ or $\mu_2(x)$), but for the data in Tables 4.45 and 4.47 (based upon $\mu_1(x)$ only) the maximum value of $\frac{n_t}{(s/\lambda)}$ is attained for $h = 0.5$ for both incidence angles.

h	n_{t}	n_{i}	A	energy balance error	$\frac{n_{\mathrm{t}}}{(s/\lambda)}$
0.025	30	30×2	30	6.7×10^{-16}	3.0
0.05	40	40×1	80	6.7×10^{-16}	4.0
0.1	50	50×2	250	1.9×10^{-15}	4.9
0.2	66	66×2	500	2.2×10^{-14}	6.0
0.25	72	72×2	550	3.8×10^{-14}	6.3
0.5	96	96×2	600	2.9×10^{-13}	6.6
1.0	128	128×2	700	3.6×10^{-13}	5.6
2.0	220	220×2	750	1.7×10^{-12}	5.3
4.0	360	360×2	800	4.6×10^{-13}	4.4

Table 4.45: Table for various h using $\mu_1(x)$ for all cases ($\frac{k}{2\pi} = 10$, $\theta = 10°$)

rep.	h	n_{t}	n_{i}	A	energy balance error	$\frac{n_{\mathrm{t}}}{(s/\lambda)}$
$\mu_2(x)$	0.025	24	24×4	500	4.4×10^{-16}	2.4
$\mu_2(x)$	0.05	36	36×4	550	1.3×10^{-15}	3.6
$\mu_1(x)$	0.1	36	36×4	600	9.8×10^{-15}	3.5
$\mu_1(x)$	0.2	72	72×2	650	1.8×10^{-15}	6.6
$\mu_1(x)$	0.25	80	80×2	650	4.8×10^{-14}	7.0
$\mu_1(x)$	0.5	110	110×2	750	1.7×10^{-14}	7.5
$\mu_1(x)$	1.0	140	140×2	850	8.4×10^{-13}	6.1
$\mu_1(x)$	2.0	220	220×2	900	3.4×10^{-13}	5.3
$\mu_1(x)$	4.0	360	360×2	1000	8.1×10^{-13}	4.4

Table 4.46: Table for various h ($\frac{k}{2\pi} = 10$, $\theta = 60°$)

h	n_{t}	n_{i}	A	energy balance error	$\frac{n_{\mathrm{t}}}{(s/\lambda)}$
0.025	24	24×4	500	4.7×10^{-15}	2.4
0.05	30	30×4	550	1.4×10^{-14}	3.0
0.1	36	36×4	600	9.8×10^{-15}	3.5
0.2	72	72×2	650	1.8×10^{-15}	6.6
0.25	80	80×2	650	4.8×10^{-14}	7.0
0.5	110	110×2	750	1.7×10^{-14}	7.5
1.0	140	140×2	850	8.4×10^{-13}	6.1
2.0	220	220×2	900	3.4×10^{-13}	5.3
4.0	360	360×2	1000	8.1×10^{-13}	4.4

Table 4.47: Table for various h using $\mu_1(x)$ for all cases ($\frac{k}{2\pi} = 10$, $\theta = 60°$)

θ	n_{fl}	n_{ch}	n_{pg}	energy balance error	time (sec)
10°	361	1025	129	4.1×10^{-14}	37
60°	361	1025	129	1.1×10^{-13}	37

Table 4.48: Results of the method of [13] for the $h = 4.0$ cases ($\frac{k}{2\pi} = 10$)

4.3.3 Dependence upon θ

Fixing $h = 0.025$, we next vary the incidence angle θ. This is done for two different wavenumbers: $k = 10 \times 2\pi$ ("resonance" regime) and $k = 1000 \times 2\pi$ ("high-frequency" regime); the incidence angles are chosen so that the wavenumbers are not a Wood Anomaly values. For $\theta = 83^\circ$ there are multiple reflections but no shadowing, while for $\theta = 87^\circ$ there are multiple reflections and shadowing (Figure 4.23). As in Section 4.3.2, the representations $\mu_1(x)$ and $\mu_2(x)$ are employed in their appropriate settings, and for purposes of further comparison we also evaluate the scattering efficiencies of the simple-reflection cases by using $\mu_1(x)$.

Certain trends in the values of the numerical parameters emerge as θ increases. For $\frac{k}{2\pi} = 10$, the number n_i of integration points per period and the integration window size A increase as θ increases, but the number n_t of target points per period slightly decreases whether $\mu_1(x)$ or $\mu_2(x)$ is used for the simple-reflection cases (Tables 4.49 and 4.50). For $\frac{k}{2\pi} = 1000$, n_t is significantly smaller for the $\mu_2(x)$-type solutions of the simple-reflection cases than it is for the $\mu_1(x)$-type solutions of the same problems and for the solutions of the other cases (Tables 4.51 and 4.52). Also, n_t for the $\mu_2(x)$-type solutions slightly increases as θ increases. However, the general trend of n_i and A increasing as θ increases still holds, and n_t for the $\mu_1(x)$-type solutions decreases as θ increases.

Due to the increasing values of n_i and A, the computational times required for these cases increase with θ. The $\frac{k}{2\pi} = 10,\ \theta = 10^\circ$ case only requires 0.17 seconds in order for its $\mu_2(x)$-type solution and scattering efficiencies to be computed, and the $\frac{k}{2\pi} = 1000,\ \theta = 10^\circ$ case takes 2.3 seconds. The $\frac{k}{2\pi} = 10,\ \theta = 87^\circ$ problem and the $\frac{k}{2\pi} = 1000,\ \theta = 89^\circ$ problem—the cases with the largest values of θ—are evaluated in 82 seconds and 456 seconds, respectively.

Remark 4.3.6. *Using the solver of [13], we accurately computed the* $\frac{k}{2\pi} = 10,\ \theta = 87^\circ$ *case in* 3.9 *seconds and the* $\frac{k}{2\pi} = 1000,\ \theta = 89^\circ$ *case in* 2078 *seconds (Table 4.53).*

4.3.4 Wood Anomaly Sensitivity

The parameter studies in the previous sections involve scattering configurations such that the wavenumbers k are well away from Wood Anomaly values—for which the periodic Green's function is undefined (Section 2.1.3.1)—except for certain cases in which k is very

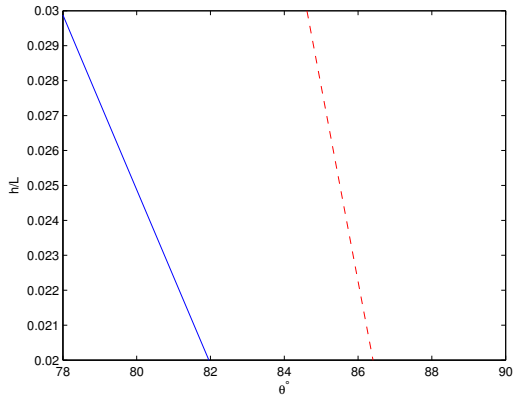


Figure 4.23: Zoomed plot of the multiple reflection threshold (solid line) and the shadowing threshold (dashed line) as a function of θ for the grating profile $f(x) = \frac{h}{2}\cos\left(\frac{2\pi x}{L}\right)$

rep.	θ	n_{t}	n_{i}	A_{sp}	A	energy balance error
$\mu_2(x)$	10°	38	38×1	0.875	30	4.4×10^{-16}
$\mu_2(x)$	60°	24	24×4	0.875	500	4.4×10^{-16}
$\mu_1(x)$	83°	24	24×5	0.875	3200	8.9×10^{-16}
$\mu_1(x)$	87°	24	24×6	0.875	11000	3.3×10^{-16}

Table 4.49: Table for various θ ($\frac{k}{2\pi} = 10$ and $h = 0.025$)

θ	n_{t}	n_{i}	A_{sp}	A	energy balance error
10°	30	30×2	0.875	30	6.7×10^{-16}
60°	24	24×4	0.875	500	4.7×10^{-15}
83°	24	24×5	0.875	3200	8.9×10^{-16}
87°	24	24×6	0.875	11000	3.3×10^{-16}

Table 4.50: Table for various θ using $\mu_1(x)$ for all cases ($\frac{k}{2\pi} = 10$ and $h = 0.025$)

rep.	θ	n_{t}	n_{i}	A_{sp}	A	energy balance error
$\mu_2(x)$	10°	16	16×200	0.175	0.2	1.8×10^{-15}
$\mu_2(x)$	60°	24	24×170	0.875	1.75	1.3×10^{-15}
$\mu_1(x)$	83°	72	72×65	0.875	350	2.0×10^{-13}
$\mu_1(x)$	87°	52	52×80	0.875	450	2.6×10^{-13}
$\mu_1(x)$	89°	52	52×90	0.875	850	2.1×10^{-13}

Table 4.51: Table for various θ ($\frac{k}{2\pi} = 1000$ and $h = 0.025$)

θ	n_{t}	n_{i}	A_{sp}	A	energy balance error
10°	260	260×10	0.175	0.2	2.3×10^{-15}
60°	160	160×25	0.875	1.75	1.3×10^{-15}
83°	72	72×65	0.875	350	2.0×10^{-13}
87°	52	52×80	0.875	450	2.6×10^{-13}
89°	52	52×90	0.875	850	2.1×10^{-13}

Table 4.52: Table for various θ using $\mu_1(x)$ for all cases ($\frac{k}{2\pi} = 1000$ and $h = 0.025$)

$\frac{k}{2\pi}$	θ	n_{fl}	n_{ch}	n_{pg}	energy balance error	time (sec)
10	87°	25	129	257	1.6×10^{-14}	3.9
1000	89°	53	8193	1025	2.0×10^{-14}	2078

Table 4.53: Results of the method of [13] ($h = 0.025$)

small or the incidence angle θ is near grazing. In this section we examine cases with wavenumbers at or very near certain Wood Anomaly values.

Remark 4.3.7. *If k is a Wood Anomaly value, then either*

$$k = k\sin(\theta) + n\frac{2\pi}{L} \longrightarrow n = \frac{kL}{2\pi}\left[1 - \sin(\theta)\right] \tag{4.4}$$

or

$$k = -k\sin(\theta) - n\frac{2\pi}{L} \longrightarrow n = -\frac{kL}{2\pi}\left[1 + \sin(\theta)\right] \tag{4.5}$$

for some integer n (Remark 2.1.5). In this section (given $L = 1$), our study of Wood Anomaly cases includes those for which $\frac{k}{2\pi}\left[1 - \sin(\theta)\right]$ and $-\frac{k}{2\pi}\left[1 + \sin(\theta)\right]$ are both integers and for which $0° \leq \theta < 90°$. Given these values of θ, the relation

$$\left|\frac{kL}{2\pi}\left[1 - \sin(\theta)\right]\right| \leq \left|-\frac{kL}{2\pi}\left[1 + \sin(\theta)\right]\right| \tag{4.6}$$

holds in this section.

As demonstrated in Appendix B, for certain problems having Wood Anomaly values for k we can use the method of this thesis to compute sets of scattering efficiencies which are accurate to machine precision and which agree with previously computed efficiencies reported in the literature. We here show that the ability to use this thesis' solver to accurately evaluate the efficiencies of other scattering configurations with Wood Anomaly wavenumbers correlates with the size of the integer $\frac{k}{2\pi}\left[1 - \sin(\theta)\right]$ for each of the cases

considered; more precisely, this ability depends upon the size of the $n = \frac{k}{2\pi}[1 - \sin(\theta)]$ Fourier mode of a given Wood Anomaly case's solution (Remark 4.3.7). Machine level accuracy for the scattering efficiencies is achieved for two Wood Anomaly cases which have somewhat large values for this integer (large enough that the size of the $n = \frac{k}{2\pi}[1 - \sin(\theta)]$ mode is very small), but only lesser accuracy is achieved for a case which has a smaller value (Section 4.3.4.1). Also, we vary k over a range of values near the wavenumber of the Wood Anomaly case which has the largest amplitude for the $n = \frac{k}{2\pi}[1 - \sin(\theta)]$ mode of its solution, and we show that in order to maintain accuracy for that set of cases the size A of the integration window must grow as k approaches the Wood Anomaly value (Section 4.3.4.2).

4.3.4.1 Computations at Wood Anomaly Values

Setting the height of the grating to be $h = 0.025$, we examine three sets of scattering problems. Each set contains a case with values of k and θ such that k is a Wood Anomaly value. Each set also contains a case with a somewhat smaller value of k and a case with a somewhat larger value of k—these k being away from all Wood Anomaly values—that together establish a performance baseline against which the computation of the Wood Anomaly case can be compared. The physical parameters for the three scattering problems with Wood Anomalies are listed in Table 4.54.

In solving these problems, the number n_t of target points per period is chosen to be large enough so that all of the significant Fourier modes of the densities are computed. This number is fixed for each set of cases. We also set the number n_i of integration points per period to be equal to n_t, since for these cases and choices of n_t this is sufficient for obtaining the most accurate solutions possible given the values of the integration window size A used.

Both $\mu_1(x)$ or $\mu_2(x)$ are computed for each of the cases. The efficiencies calculated using the $\mu_2(x)$ solutions are tabulated (the $\mu_1(x)$-based efficiencies are essentially identical), and the Fourier amplitudes of both the $\mu_1(x)$ and $\mu_2(x)$ solutions are plotted.

We are able to use the method of this thesis to compute the efficiencies of all of the cases in Set 1—for which $\frac{k}{2\pi}$ is at or around 10 and $\theta = 0°$—to machine precision (Table 4.55). No difference between the cases in the convergence of their efficiencies in A is observed. Additionally, as shown in Figure 4.24, the Fourier amplitudes of the solution (either $\mu_1(x)$ or $\mu_2(x)$) of the $\frac{k}{2\pi} = 9.5$ case are similar to those of the $\frac{k}{2\pi} = 10.5$ case's solution, and most

of the amplitudes of the $\frac{k}{2\pi} = 10$ case's solution (i.e., the solution of the Wood Anomaly case) are also similar. But, the $n = -10$ and $n = 10$ modes' amplitudes of this case's solution, which are approximately 10^{-11} in magnitude, differ from those of the other cases' solutions; these values of n correspond to the values $-\frac{k}{2\pi}[1 + \sin(\theta)] = -10$ and $\frac{k}{2\pi}[1 - \sin(\theta)] = 10$ listed in Table 4.55 for the $\frac{k}{2\pi} = 10$ case in this set.

The cases of Set 2 have the same values for the wavenumbers k but an increased incidence angle θ ($\theta = 30°$) relative to the Set 1 cases, and—unlike for Set 1—the efficiencies for the Wood Anomaly case in this set cannot be computed to machine precision (Table 4.56). They can be computed to a similar accuracy as the efficiencies of the other cases in the set when using $A = 5$ (the energy balance error levels are about 10^{-7}–10^{-6} for all three cases). But, their accuracy is significantly worse than the accuracies of the efficiencies of the other cases when using $A = 200$ (the energy balance error level is about 10^{-9} rather than about 10^{-15}), and this accuracy does not dramatically improve even when increasing A to 50000. The amplitude of the $n = \frac{k}{2\pi}[1 - \sin(\theta)] = 5$ mode of the solution of the Wood Anomaly case substantially differs from the $n = 5$ modes of the other cases' solutions (Figure 4.25); this amplitude is approximately 10^{-5} in magnitude, which is much larger than the $n = -10$ and $n = 10$ modes' amplitudes of the Set 1 Wood Anomaly case's solution. The $n = -15$ modes of all three cases ($-\frac{k}{2\pi}[1 + \sin(\theta)] = -15$ for the Wood Anomaly case) are too small in magnitude to be numerically significant, however.

The incidence angle $\theta = 30°$ is maintained for the Set 3 cases, but $\frac{k}{2\pi}$ is increased to be at or around 100. We find that machine-level accuracies can once again be achieved for all three cases (Table 4.57), just as they were obtained for the Set 1 cases. For the Wood Anomaly case in this set, $-\frac{k}{2\pi}[1 + \sin(\theta)] = -150$ and $\frac{k}{2\pi}[1 - \sin(\theta)] = 50$; the $n = -150$ and $n = 50$ modes of all three cases are numerically insignificant (Figure 4.26), and thus none of the Fourier amplitudes of the Wood Anomaly case noticeably differ from the corresponding ones of the other cases in this set.

We conclude, therefore, that our ability to accurately evaluate the scattering efficiencies of these Wood Anomaly cases by using the method of this thesis depends upon the size of the quantity $\frac{k}{2\pi}[1 - \sin(\theta)]$, particularly as it bears upon the sizes of the $n = \frac{k}{2\pi}[1 - \sin(\theta)]$ modes of their solutions. The solutions of the Wood Anomaly cases for which the $n = \frac{k}{2\pi}[1 - \sin(\theta)]$ modes are very small (10^{-11} or less in magnitude) can be computed sufficiently accurately so that the efficiencies have an energy balance error indicative of machine-

Set of cases	$\frac{k}{2\pi}$	θ	$-\frac{k}{2\pi}\left[1+\sin(\theta)\right]$	$\frac{k}{2\pi}\left[1-\sin(\theta)\right]$
1	10	0°	-10	10
2	10	30°	-15	5
3	100	30°	-150	50

Table 4.54: The physical parameters for each Wood Anomaly case ($h = 0.025$)

$\frac{k}{2\pi}$	A	energy balance error
9.5	5	1.4×10^{-8}
10	5	9.8×10^{-9}
10.5	5	5.7×10^{-9}
9.5	30	8.9×10^{-16}
10	30	3.3×10^{-16}
10.5	30	8.9×10^{-16}

Table 4.55: Table for the Set 1 cases, computed using $\mu_2(x)$ ($h = 0.025$, $\theta = 0^\circ$, $n_\mathrm{t} = 48$, $n_\mathrm{i} = 48 \times 1$)

level accuracy, but for the Set 2 Wood Anomaly case (for which the $n = \frac{k}{2\pi}\left[1-\sin(\theta)\right]$ mode is about 10^{-5} in magnitude) only about single precision accuracy is achieved.

4.3.4.2 Computations as k Approaches Wood Anomaly Values

As evidenced by the Set 1 and Set 3 cases in the previous section, our numerical computations of the efficiencies of certain Wood Anomaly cases converge to machine-level accuracy just as rapidly in the integration window size A as they do for cases with similar, but not Wood Anomaly-valued, wavenumbers (Tables 4.55 and 4.57). Since this is not true for the cases of Set 2, however, we fix $h = 0.025$ and $\theta = 30^\circ$ (just as they are fixed for the Set 2 cases) and compute the efficiencies of scattering cases with wavenumbers k approaching

$\frac{k}{2\pi}$	A	energy balance error
9.5	5	1.2×10^{-6}
10	5	2.0×10^{-7}
10.5	5	4.2×10^{-7}
9.5	200	1.2×10^{-15}
10	200	2.7×10^{-9}
10	50000	1.3×10^{-10}
10.5	200	1.9×10^{-15}

Table 4.56: Table for the Set 2 cases, computed using $\mu_2(x)$ ($h = 0.025$, $\theta = 30^\circ$, $n_\mathrm{t} = 60$, $n_\mathrm{i} = 60 \times 1$)

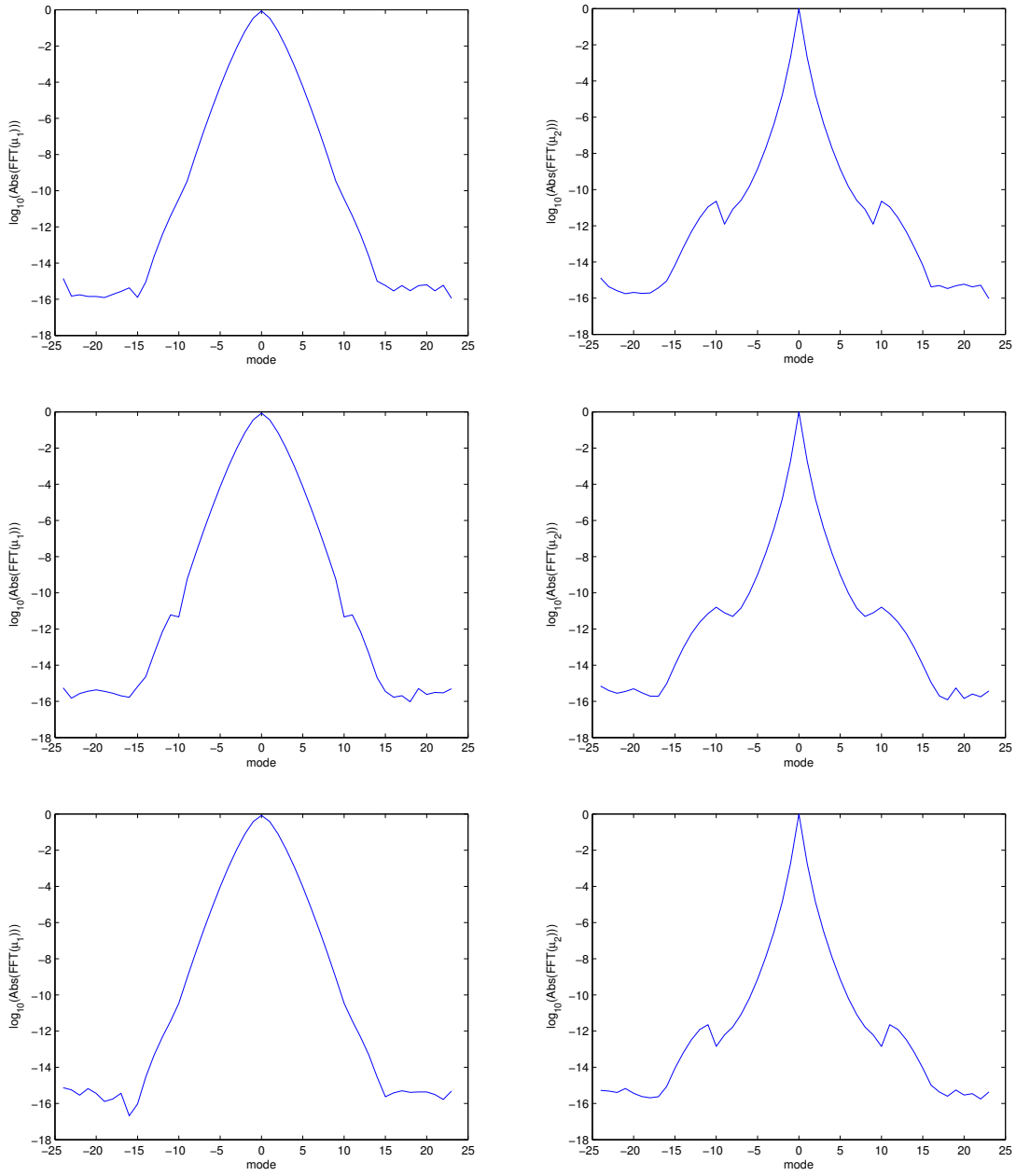


Figure 4.24: Set 1: Fourier amplitudes of $\mu_1(x)$ (left) and $\mu_2(x)$ (right) for $\frac{k}{2\pi} = 9.5$ (top), $\frac{k}{2\pi} = 10$ (middle) and $\frac{k}{2\pi} = 10.5$ (bottom)

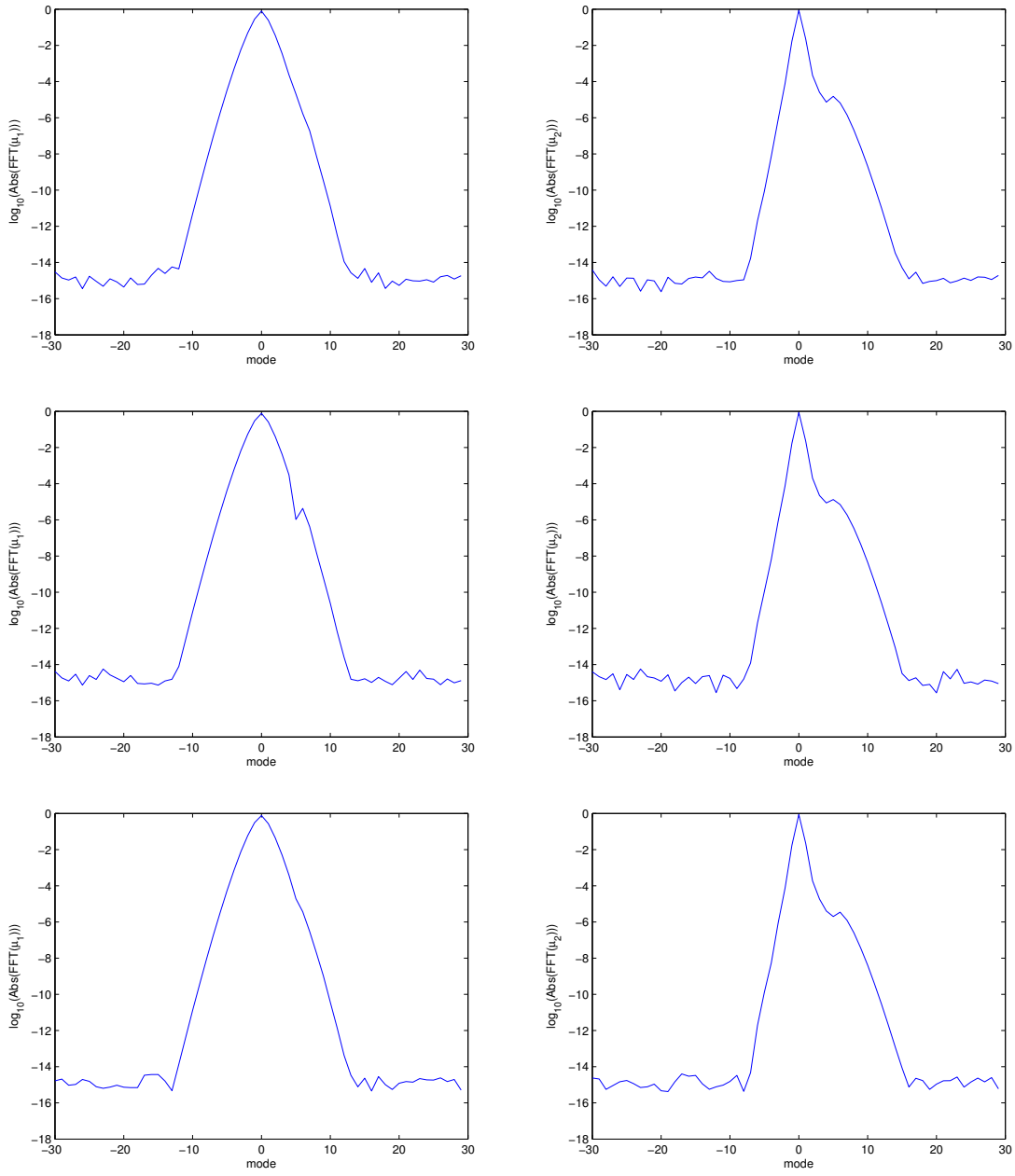


Figure 4.25: Set 2: Fourier amplitudes of $\mu_1(x)$ (left) and $\mu_2(x)$ (right) for $\frac{k}{2\pi} = 9.5$ (top), $\frac{k}{2\pi} = 10$ (middle) and $\frac{k}{2\pi} = 10.5$ (bottom)

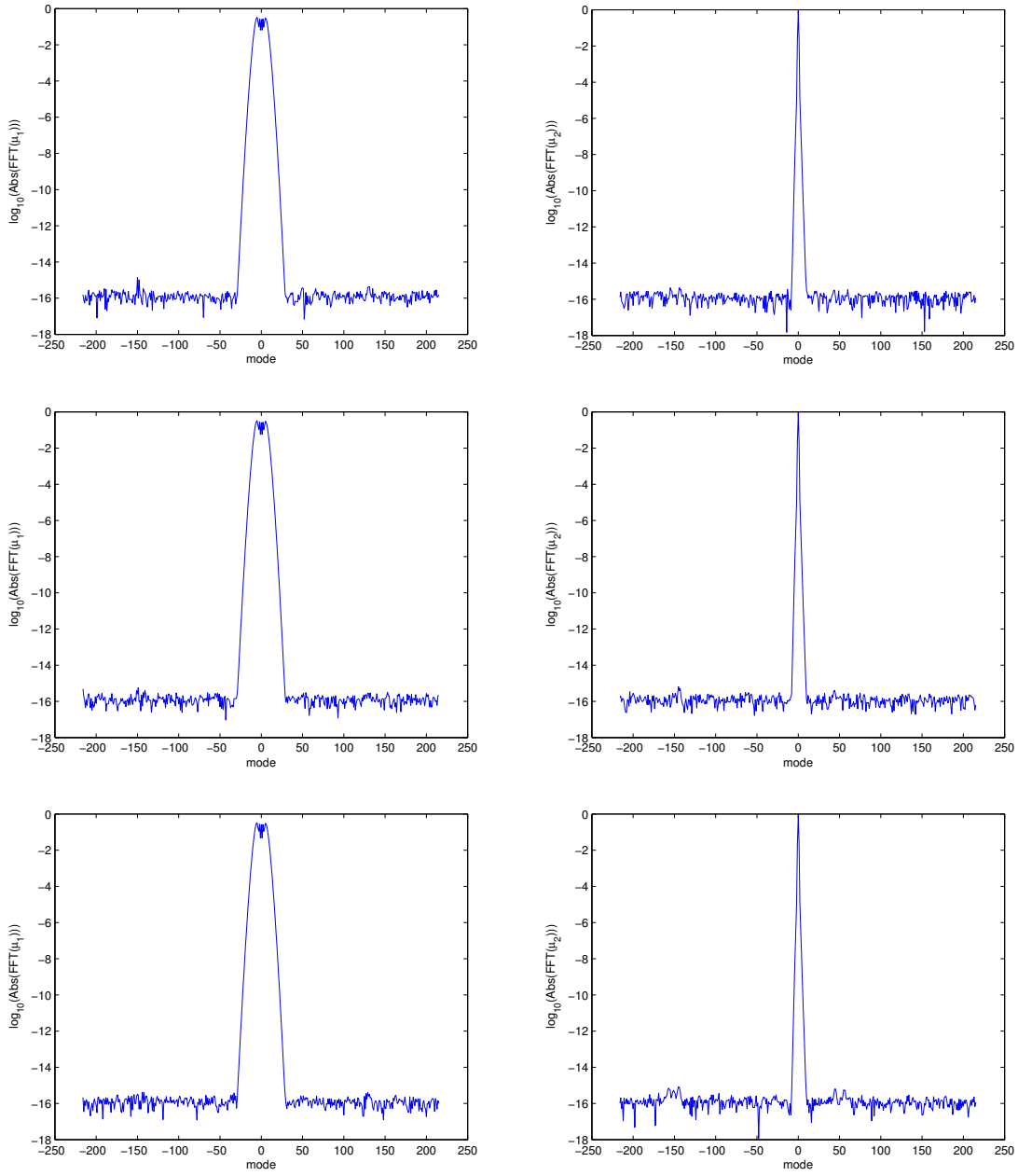


Figure 4.26: Set 3: Fourier amplitudes of $\mu_1(x)$ (left) and $\mu_2(x)$ (right) for $\frac{k}{2\pi} = 99.5$ (top), $\frac{k}{2\pi} = 100$ (middle) and $\frac{k}{2\pi} = 100.5$ (bottom)

$\frac{k}{2\pi}$	A	energy balance error
99.5	1	5.5×10^{-9}
100	1	1.2×10^{-8}
100.5	1	5.9×10^{-9}
99.5	4	2.4×10^{-15}
100	4	2.9×10^{-15}
100.5	4	2.0×10^{-15}

Table 4.57: Table for the Set 3 cases, computed using $\mu_2(x)$ ($h = 0.025$, $\theta = 30°$, $n_t = 432$, $n_i = 432 \times 1$)

$\frac{k}{2\pi}$	A	energy balance error	$\left\|5 - \frac{k}{2\pi}\left[1 - \sin(\theta)\right]\right\|$	$\left\|5 - \frac{k}{2\pi}\left[1 - \sin(\theta)\right]\right\| A$
9.5	100	5.0×10^{-14}	0.25	25
9.75	200	7.4×10^{-14}	0.125	25
9.9	500	8.1×10^{-14}	0.05	25
9.99	5000	8.5×10^{-14}	0.005	25
9.999	50000	9.2×10^{-14}	0.0005	25
10.001	50000	8.9×10^{-13}	0.0005	25
10.01	5000	1.6×10^{-13}	0.005	25
10.1	500	6.9×10^{-14}	0.05	25
10.25	200	6.9×10^{-14}	0.125	25
10.5	100	7.3×10^{-14}	0.25	25

Table 4.58: Table for various k ($h = 0.025$, $\theta = 30°$, $n_t = 60$, $n_i = 60 \times 1$)

that of the Wood Anomaly case in Set 2 (for which $\frac{k}{2\pi} = 10$ and $\frac{k}{2\pi}\left[1 - \sin(\theta)\right] = 5$).

Given the terms $\frac{1}{\{kL[1+\sin(\theta)]+2\pi s\}\frac{cA}{L}}$, $s \in \mathbb{Z}$ and $\frac{1}{\{kL[1-\sin(\theta)]+2\pi s\}\frac{cA}{L}}$, $s \in \mathbb{Z}$ that arise from integration by parts in the error estimates of Theorem 3.1.9—see Section 3.1.3.3, in particular (3.179)—we anticipate that A must increase as $\frac{k}{2\pi}\left[1 - \sin(\theta)\right]$ approaches 5 in order to maintain computational accuracy. Table 4.58 confirms this property: the energy balance errors achieved over the intervals $k \in [9.5, 9.999]$ and $k \in [10.001, 10.5]$ are nearly constant if $\left|5 - \frac{k}{2\pi}\left[1 - \sin(\theta)\right]\right| A$ is kept constant. Table 4.59 provides further confirmation by showing how leaving A fixed at 100 as $\frac{k}{2\pi}$ increases from 9.5 to 9.75 results in a significant drop in overall accuracy, while increasing A to 400 rather than to 200 leads to a noticeable improvement in overall accuracy.

$\frac{k}{2\pi}$	A	energy balance error	$5-\frac{k}{2\pi}\left[1-\sin(\theta)\right]$	$\left\{5-\frac{k}{2\pi}\left[1-\sin(\theta)\right]\right\}A$
9.5	100	5.0×10^{-14}	0.25	25
9.75	100	1.5×10^{-12}	0.125	12.5
9.75	200	7.4×10^{-14}	0.125	25
9.75	400	1.8×10^{-15}	0.125	50

Table 4.59: Table for various k ($h=0.025$, $\theta=30^\circ$, $n_\mathrm{t}=60$, $n_\mathrm{i}=60\times1$)

Chapter 5

Conclusions and Future Work

5.1 Conclusions

We have presented a new integral equation method for the solution of direct rough surface scattering problems in two dimensions. This approach—motivated, in part, by the bounded obstacle solver introduced in [12]—was described in detail in Chapter 2. Through the use of smooth windowing functions, we created an efficient alternative to the calculation of the periodic Green's function that is typically employed in other integral equation methods for these problems. We derived two different analytical representations for the solutions; after an examination of the types of scattering that can occur and the conditions which give rise to these phenomena, we showed how the use of one or the other of these representations is advantageous depending upon whether multiple scattering is present. Combining the windowing and the two representations, we generated certain approximating scattering equations that are solved numerically using a spectral quadrature rule, FFTs and GMRES—allowing for a smaller number n_{t} of target points for the solutions than the number n_{i} of quadrature points in order to further significantly enhance performance in many high-frequency cases.

Chapter 3 included many proofs regarding this solver. We demonstrated the convergence of the integral operators of the approximating integral equations to those of the full integral equations as the window size A becomes infinite, and we showed how this implies the existence and convergence of the approximating solutions to the full solutions. In particular, we established that these convergences are super-algebraic in A as $A \to \infty$ for scattering configurations with C^∞ surface profiles. Similarly, we proved the convergence of the quadrature rule as $n_{\mathrm{i}} \to \infty$; this convergence was shown to be super-algebraic in n_{i} as

$n_{\mathrm{i}} \to \infty$ for problems with smooth gratings, and corresponding existence and convergence proofs for the numerical solutions were also given. Finally, we argued that the method requires $\mathcal{O}(1)$ computational times as the wavenumber k increases without bound (for a given scattering profile and incidence angle) in order to compute the solutions of simple-reflection cases to a fixed level of accuracy.

In Chapter 4, we provided a large number of numerical results which demonstrated the accuracy and efficiency of the solver under a variety of scattering configurations. We confirmed computationally the convergence results of the previous chapter. Also, we showed that the performance of our method compares very favorably to that of other leading integral equation approaches as well as the Kirchhoff approximation for many wavenumbers and scattering surfaces (including sinusoidal, multi-scale and simulated ocean surfaces). Finally, we studied how the computational parameters of the solver need to be adjusted as certain physical quantities are systematically varied; in particular, $\mathcal{O}(1)$ in k computational times indeed were achieved as k became large (for certain scattering profiles and incidence angles) for simple-reflection cases, a maximum in the numbers of degrees of freedom per wavelength for the solutions was observed as h was increased (all other physical parameters being kept constant), certain cases with Wood Anomaly values for k were computed to machine precision and computational accuracy was maintained for another set of cases (with a fixed scattering profile and incidence angle) by increasing A in an anticipated manner (given the proofs of the previous chapter) as k approached a Wood Anomaly value.

5.2 Future Work

There is much potentially fruitful work that could be undertaken in the future given the research presented in this thesis. We list some of the possibilities here. They include the application of our solver to additional scattering configurations of interest. Also, various improvements and extensions of the algorithm—such as the acceleration of certain calculations and a three-dimensional version of the method—are noted.

Solutions of rough surface scattering problems with high-frequency configurations, deep grating profile configurations, low grazing angle configurations, etc., are rapidly computed by the approach of this thesis—thus potentially facilitating further examination of cases that are of scientific and engineering interest. Better simulations of phenomena such as the

scattering of acoustic and electromagnetic waves from random surfaces, for example, could be pursued using this solver. Also, the smooth windowing functions that we employ may allow our approach to be applied profitably to cases containing finite (rather than periodic) rough surfaces, just as numerical tapering was successfully used for such problems in [60]. In addition, due to its excellent accuracy and efficiency, our algorithm could form a key building block in the construction of a powerful method for solving inverse problems.

There also are a number of improvements and extensions that could be made to our solver. For example, a straightforward quadrature is currently employed for calculating each scattering efficiency, which implies that in a high-frequency case the determination of the efficiencies dominates the computational time. It may be possible to construct an asymptotic approximation to the efficiencies' formula so that accurate values could be computed significantly more rapidly when k is large. Also, other quantities related to the scattered field (e.g., its magnitude at various points near the scattering surface) could prove to be useful alternatives to the efficiencies, depending upon the application.

Additionally, the use of FFT acceleration techniques may allow for further reduction in computational times. A particular accelerator for three-dimensional problems is described in [14, 15]; it could be suitably modified to treat the configurations we have considered in this thesis.

One of the most important extensions of the method of this thesis would include, of course, an approach to solving three-dimensional scattering problems containing two-dimensional periodic rough surfaces that vary in z. The techniques underlying the solver discussed in [14, 15]—including the use of partitions of unity, analytical resolution of singularities and the aforementioned FFT acceleration—should prove useful towards this end. Some of these techniques already have been successfully incorporated in the three-dimensional method introduced in [11], which is an extension of the two-dimensional bounded obstacle solver that was originally presented in [12] and is closely related to our work here.

The two-dimensional solver of [12]—described in Appendix A of this thesis—includes unique methods for treating the multiple reflection phenomena that arise in cases with non-convex scatterers as well as the shadowing phenomena that always occur. For the configuration in [12] which gives rise to multiple reflections, a modified ansatz for the density that accounts for all of the directions of propagation of the reflected geometrical optics rays is employed. This approach also could be made to work for rough surface problems

containing multiple reflections. Likewise, the change of variables used by the bounded obstacle algorithm near shadow boundaries could be suitably adapted; alternatively, it may be possible to construct an ansatz for the solution that accounts for both the multiple reflections and the creeping waves that arise in cases with shadowing.

The scattering profiles that were considered in this thesis are at least twice continuously differentiable. Less differentiable surfaces, such as those containing corners, give rise to additional scattering phenomena beyond what we have discussed (see [7] for details). Changes of variables (such as those used in [20, 30]) would allow for high-order numerical convergence (in the number of discretization points) to be preserved for configurations with such surfaces; a modified ansatz for the density that accounts for the additional types of scattering that occur may be required in order to maintain rapid computational times for high-frequency cases.

Another possible modification of our method is to use integral equations based upon half space Green's functions rather than the ones we developed via the free space Green's function. The Nyström method of [44], for example, uses the half space Green's function for the Helmholtz equation with a Dirichlet boundary condition; the "quasi-periodic Dirichlet Green's function" that is employed in the approaches of [4] is related to this half space Green's function in the same way that the periodic Green's function is related to the free space Green's function. For points $\boldsymbol{r}, \boldsymbol{r}'$ on the scattering surface, the half space Green's function in [44] decays more rapidly than the free space Green's function does as $|\boldsymbol{r} - \boldsymbol{r}'| \to \infty$, and, unlike the periodic Green's function, the quasi-periodic Dirichlet Green's function in [4] can be defined even for Wood Anomaly wavenumbers. Thus, it is of interest to determine if the substitution of half space Green's functions into our approach results in significant performance improvements in general (e.g., smaller integration window sizes being required to achieve certain accuracy levels) and any noticeably different results for Wood Anomaly cases in particular.

Additionally, the benefits of our approach need not be limited to problems with perfectly reflecting scattering surfaces. Transmission of incident electromagnetic waves occurs in cases with periodic interfaces between two dielectric materials (e.g., air and ocean water), and this gives rise to coupled systems of integral equations (as described, for example, in [16]). The integral operators within such systems are closely related to the ones that were described in this thesis, and we therefore anticipate that our algorithm could be adapted to solve

dielectric rough surface problems.

Appendix A

A treatment of bounded obstacles

This work originally was presented in [12]. Included are some minor changes of formatting and content relative to that paper.

A.1 Introduction

However efficient, direct numerical methods for the solution of scattering problems require a fixed number of discretization points per wavelength λ, and thus exhibit a computational complexity of at least $\mathcal{O}(k^n)$ for an n-dimensional discretization (where $k = 2\pi/\lambda$ is the wave number). It is therefore desirable to produce numerical methods which remain efficient as the frequency (and, thus, the size of the problem) grows. If accurate high-frequency solvers are made available with a bounded computational complexity as the frequency tends to infinity (that is, methods with an asymptotic $\mathcal{O}(1)$ computational complexity), then one can envision the development of a computational capability allowing the solution of essentially arbitrary scattering problems.

This appendix presents such an $\mathcal{O}(1)$ solver for surface-scattering problems by convex obstacles in two or three dimensions, using a combined-field integral equation [41]. Our rigorous (convergent) approach relies on two main elements [9, 10].

The first of these elements is a transformation of a boundary integral equation which allows it to explicitly capture, with coarse discretizations, the rapidly oscillatory progression of the surface currents. For this purpose, an ansatz derived from asymptotic theory [34] is used: the original unknown in the boundary integral formulation is replaced by the product of a slowly varying amplitude and a highly-oscillatory exponential; see Section A.2.1. The slowly varying amplitude can then be represented by a number of degrees of freedom

independent of the frequency. This idea is similar to those presented by [38] and [29] for partial differential equations, and by [31], [2], [1] and [47] for integral equations. Unlike the previous approaches, however, the present treatment accounts rigorously for the fact that the ansatz is only valid in certain regions of the scattering surface.

The second main element in the present algorithm is a localized integration method related to the method of stationary phase. This localized integration scheme, which reduces the support of integration to a small subset of the scattering surface, can be seen as a natural link between high-frequency approximate, *non-convergent* methods such as the Kirchhoff approximation, and a direct integral equation method. As discussed below, the size of the reduced integration support is related to the wavelength, leading to a number of integration points independent of frequency, and thus, to a frequency-independent overall computational complexity.

In addition to these main elements, our solver uses high-order discretization schemes for accuracy: the Nyström method described in [20] in two dimensions, and the method described in [14, 15] in three dimensions. In all cases, the high-order nature of the high-frequency solver is achieved through use of Fourier interpolation and the trapezoidal rule for integration of periodic functions: see Section A.4.

The numerical method is then completed through use of a matrix-free Krylov subspace linear algebra solver. The result is a high-order *convergent* algorithm that can solve accurately scattering problems throughout the electromagnetic spectrum, and can deliver error-controllable solutions in computational times that are independent of frequency. We illustrate the efficiency of this algorithm through a series of computational results in Section A.5; in particular, we demonstrate the high-order convergence of the solver as well as its asymptotically bounded computational complexity as the frequency increases: see Table A.3. The extension of the method to *non-convex* scatterers is finally discussed briefly in Section A.6.

A.2 Boundary Integral Formulation

We consider the problem of evaluating the scattering of an incident plane wave $\psi^{\text{inc}}(\boldsymbol{r}) = e^{ik\boldsymbol{\alpha}\cdot\boldsymbol{r}}$, $|\boldsymbol{\alpha}| = 1$, from a convex impenetrable obstacle D. We thus look for the solution

$\psi(\boldsymbol{r}) = \psi^{\text{inc}}(\boldsymbol{r}) + \psi^{\text{scat}}(\boldsymbol{r})$ of the Helmholtz equation under Dirichlet boundary conditions

$$\Delta\psi(\boldsymbol{r}) + k^2\psi(\boldsymbol{r}) = 0 \quad \text{in} \quad \mathbb{R}^n\backslash\bar{D}, \quad n = 2 \text{ or } 3, \tag{A.1}$$

$$\psi = 0 \quad \text{on} \quad \partial D, \tag{A.2}$$

where the scattered field $\psi^{\text{scat}}(\boldsymbol{r})$ satisfies the Sommerfeld radiation condition [20]. For the sake of simplicity we treat a scalar scattering problem—acoustic or TE-electromagnetic; the full electromagnetic problem can be handled in a similar way.

A.2.1 Ansatz

To introduce some of the issues arising in our high-frequency integral method, let us consider the following boundary integral formulation of the problem (A.1)–(A.2), which takes as the unknown function the boundary values of the normal derivative:

$$\begin{aligned}\frac{1}{2}\frac{\partial\psi(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})} = \left(\frac{\partial\psi^{\text{inc}}(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})} + i\gamma\psi^{\text{inc}}(\boldsymbol{r})\right) - \int_{\partial D}\frac{\partial\Phi(\boldsymbol{r},\boldsymbol{r}')}{\partial\boldsymbol{\nu}(\boldsymbol{r})}\frac{\partial\psi(\boldsymbol{r}')}{\partial\boldsymbol{\nu}(\boldsymbol{r}')}\,ds(\boldsymbol{r}') \\ - i\gamma\int_{\partial D}\Phi(\boldsymbol{r},\boldsymbol{r}')\frac{\partial\psi(\boldsymbol{r}')}{\partial\boldsymbol{\nu}(\boldsymbol{r}')}\,ds(\boldsymbol{r}'),\end{aligned} \tag{A.3}$$

where $\boldsymbol{\nu}(\boldsymbol{r})$ is the external normal to the surface at point $\boldsymbol{r}$ and where $\Phi(\boldsymbol{r},\boldsymbol{r}')$ equals $i/4\,H_0^{(1)}(k|\boldsymbol{r}-\boldsymbol{r}'|)$ in two dimensions and $e^{ik|\boldsymbol{r}-\boldsymbol{r}'|}/(4\pi|\boldsymbol{r}-\boldsymbol{r}'|)$ in three dimensions. In this equation, γ is an arbitrary positive constant. Following [14, 15] we use $\gamma = \max\{3, D/\lambda\}$ (where D is the diameter of the scatterer), which gives rise to rapid convergence of the linear algebra iterative solver.

As mentioned above, our high-frequency approach is based on a high-frequency ansatz for the unknown

$$\mu(\boldsymbol{r}) = \frac{\partial\psi(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})} \tag{A.4}$$

of the problem. For a convex scatterer, our ansatz reads

$$\mu(\boldsymbol{r}) = \mu_{\text{slow}}(\boldsymbol{r})\,e^{ik\boldsymbol{\alpha}\cdot\boldsymbol{r}}, \tag{A.5}$$

where the new unknown μ_{slow} is assumed to be a slowly oscillatory function of $\boldsymbol{r} \in \partial D$; see Section A.2.2 and Section A.3.3 for details. The validity of (A.5) in a portion of the scattering surface indicates that, on that portion, the unknown μ oscillates along with

the incident field. For non-convex scatterers (or, more generally, in presence of multiple reflections), a more elaborate ansatz can be constructed using ray-tracing (GO) techniques; see Section A.6.

As it happens, only the solution of certain types of integral equations can be represented through an ansatz of this type. As a rule, an integral equation whose unknown is a *physical quantity* can be represented by an ansatz of this form—the unknown in (A.3) is the normal derivative of the solution, and it therefore admits such a representation. In contrast, the density φ in the integral equation [20]

$$\frac{1}{2}\varphi(\boldsymbol{r}) = \psi^{\text{inc}}(\boldsymbol{r}) - \int_{\partial D} \frac{\partial \Phi(\boldsymbol{r},\boldsymbol{r}')}{\partial \boldsymbol{\nu}(\boldsymbol{r}')}\varphi(\boldsymbol{r}')\, ds(\boldsymbol{r}') + i\gamma \int_{\partial D} \Phi(\boldsymbol{r},\boldsymbol{r}')\varphi(\boldsymbol{r}')\, ds(\boldsymbol{r}') \tag{A.6}$$

for our Dirichlet problem (A.1)–(A.2) does not admit such a representation (see Figure A.1). The question does naturally arise: What is the difference in character between the integral formulations (A.3) and (A.6)?

This can be understood through the consideration of a simple scattering surface: a pair of parallel planes. It is easy to check that the the combination of integrals in (A.3) *integrated over the illuminated plane only* produces field values on the non-illuminated surface which equal, precisely, the value of the inhomogeneous term in (A.3) on the non-illuminated boundary. It follows that the unknown function vanishes on the non-illuminated boundary, and therefore the integral over that boundary does not give rise to additional fields on the illuminated boundary. Thus, a solution of the equation can be obtained, in this case, by consideration of scattering by the illuminated surface alone. This is not true for equation (A.6). Indeed, in this case further corrections on the illuminated surface must be introduced, as the non-illuminated surface 'scatters' a field into the illuminated surface, which then gives rise to additional fields on the non-illuminated surface, and so on—so that use of the expression (A.5) in conjunction with equation (A.6) results in a highly oscillatory μ_{slow}. Considerations related to these can be used to determine whether, for general, non-planar surfaces, the solutions of a given integral equation satisfy an ansatz of the form (A.5). Indeed, while such a discussion would generally not be exact for finite wave numbers and curved surfaces, these arguments can be used *asymptotically* as $k \to \infty$—which suffices to determine the validity (or lack of validity) of our integral ansatz for a given integral equation.

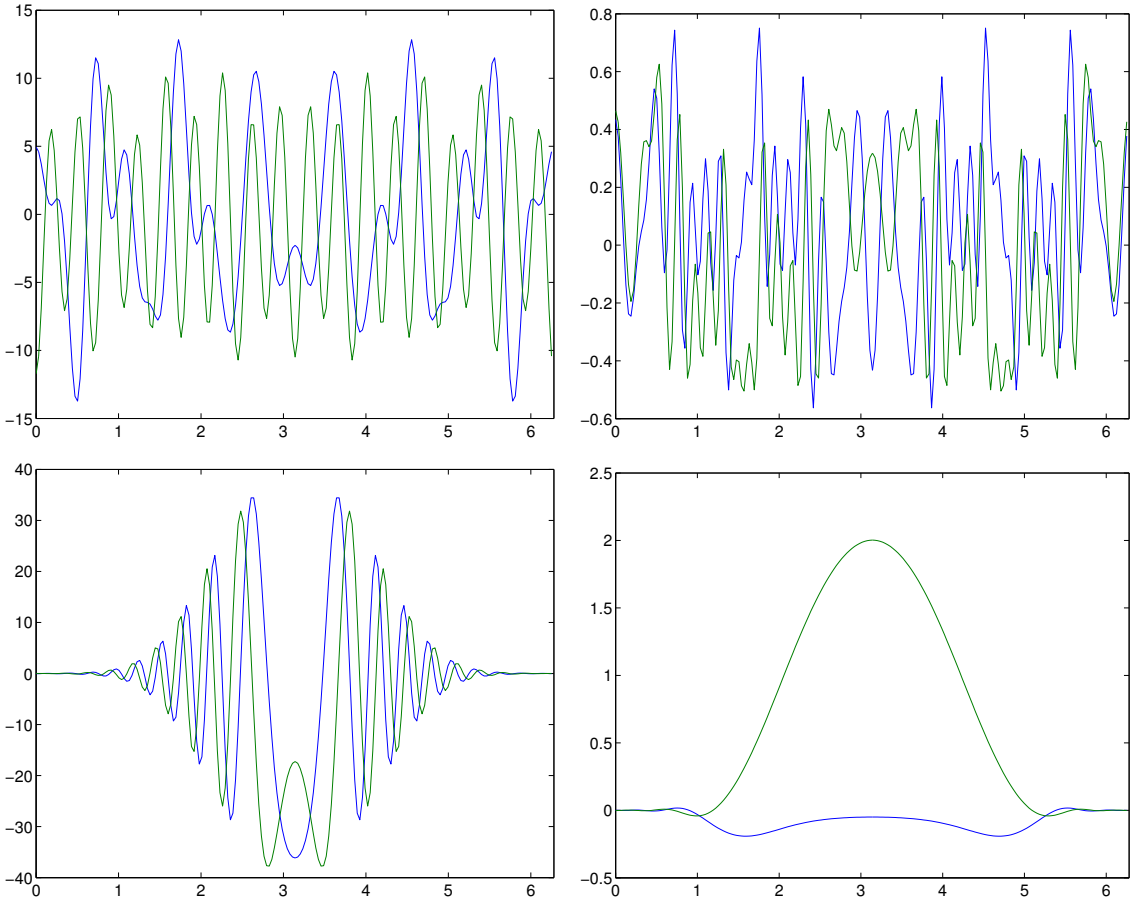


Figure A.1: Scattering by a circular cylinder $\boldsymbol{r}(\theta) = (a\cos\theta, a\sin\theta)$, $0 \le \theta \le 2\pi$, with $ka = 20$; all graphs show real and imaginary parts of complex quantities as functions of the angular coordinate θ. Top left: $\varphi(\boldsymbol{r}(\theta))$ (solution of (A.6)). Top right: $\varphi(\boldsymbol{r}(\theta))/(ke^{ik\boldsymbol{\alpha}\cdot\boldsymbol{r}(\theta)})$. Bottom left: $\mu(\boldsymbol{r}(\theta)) = \frac{\partial\Psi(\boldsymbol{r}(\theta))}{\partial\boldsymbol{\nu}(\boldsymbol{r}(\theta))}$ (solution of (A.3)). Bottom right: $\mu(\boldsymbol{r}(\theta))/(ke^{ik\boldsymbol{\alpha}\cdot\boldsymbol{r}(\theta)})$.

A.2.2 High-Frequency Integral Equation

Using (A.4), the boundary integral formulation (A.3) can be rewritten as

$$\frac{1}{2}\mu(\boldsymbol{r}) + (K'\mu)(\boldsymbol{r}) + i\gamma(S\mu)(\boldsymbol{r}) = \frac{\partial\psi^{\text{inc}}(\boldsymbol{r})}{\partial\boldsymbol{\nu}(\boldsymbol{r})} + i\gamma\psi^{\text{inc}}(\boldsymbol{r}), \quad \boldsymbol{r}\in\partial D, \tag{A.7}$$

with

$$(S\mu)(\boldsymbol{r}) = \int_{\partial D}\Phi(\boldsymbol{r},\boldsymbol{r}')\mu(\boldsymbol{r}')\,ds(\boldsymbol{r}'), \tag{A.8}$$

$$(K'\mu)(\boldsymbol{r}) = \int_{\partial D}\frac{\partial\Phi(\boldsymbol{r},\boldsymbol{r}')}{\partial\boldsymbol{\nu}(\boldsymbol{r})}\mu(\boldsymbol{r}')\,ds(\boldsymbol{r}'). \tag{A.9}$$

Introducing the ansatz (A.5) in (A.7) and dividing by $e^{ik\boldsymbol{\alpha}\cdot\boldsymbol{r}}$, we obtain

$$\frac{1}{2}\mu_{\text{slow}}(\boldsymbol{r}) + (\tilde{K}'\mu_{\text{slow}})(\boldsymbol{r}) + i\gamma(\tilde{S}\mu_{\text{slow}})(\boldsymbol{r}) = i\boldsymbol{\nu}\cdot\boldsymbol{\alpha}k + i\gamma, \quad \boldsymbol{r}\in\partial D, \tag{A.10}$$

where $\tilde{S}$ and $\tilde{K}'$ denote the integral operators

$$(\tilde{S}\mu_{\text{slow}})(\boldsymbol{r}) = \int_{\partial D}\Phi(\boldsymbol{r},\boldsymbol{r}')e^{ik\boldsymbol{\alpha}\cdot(\boldsymbol{r}'-\boldsymbol{r})}\mu_{\text{slow}}(\boldsymbol{r}')\,ds(\boldsymbol{r}'), \tag{A.11}$$

$$(\tilde{K}'\mu_{\text{slow}})(\boldsymbol{r}) = \int_{\partial D}\frac{\partial\Phi(\boldsymbol{r},\boldsymbol{r}')}{\partial\boldsymbol{\nu}(\boldsymbol{r})}e^{ik\boldsymbol{\alpha}\cdot(\boldsymbol{r}'-\boldsymbol{r})}\mu_{\text{slow}}(\boldsymbol{r}')\,ds(\boldsymbol{r}'). \tag{A.12}$$

As we shall show, except for certain oscillatory behavior of μ_{slow} at the shadow boundaries (see Section A.3.3), the kernels in equations (A.11) and (A.12) are the only highly-oscillatory functions in the boundary integral formulation (A.10). Since μ_{slow} is a slowly varying function away from the shadow boundaries, this density can be represented, to within any prescribed tolerance, by a *fixed* set of discretization points, independent of frequency.

A.3 Localized Integration

Despite the fact that the unknown in the modified boundary integral formulation (A.10) is a slowly oscillating function, a direct numerical evaluation of the integrals (A.11) and (A.12) would still require a number of quadrature points proportional to the wave number k. In this section we show that an extension of the method of stationary phase [5] can be

used to produce a *convergent* method which requires a fixed number of quadrature points for prescribed accuracies and arbitrary wave numbers.

A.3.1 Stationary Phase

To incorporate ideas implicit in the method of stationary phase we first obtain the critical points of the integrals in (A.11) and (A.12). The details of such an evaluation depend on the particular kernels under consideration, but in the present case, for $\boldsymbol{r} \neq \boldsymbol{r'}$, both kernels in (A.11) and (A.12) behave asymptotically as

$$e^{ik[|\boldsymbol{r}-\boldsymbol{r}'|+\boldsymbol{\alpha}\cdot(\boldsymbol{r}'-\boldsymbol{r})]} = e^{ik\phi}, \tag{A.13}$$

i.e., as the kernel of a generalized Fourier integral with phase ϕ. The critical points are thus

1. the target (observation) point $\boldsymbol{r}$ itself, where the kernel is singular;

2. the stationary points, i.e., the points where the phase ϕ in the integrals has a vanishing gradient. (Note that these stationary points vary as a function of the target point, and that both the first and second derivatives of the phase vanish at the shadow boundaries.)

In Section A.8 below we present, as an example, the details of the evaluation of the corresponding stationary points for a TE integral equation.

In view of the method of stationary phase we know that, *asymptotically*, the only significant contributions to the integrals (A.11) and (A.12) arise from values of the slow integrands and their derivatives at the critical points. In order to construct a *convergent* method for arbitrary frequencies, we introduce an integration procedure based on localization *around* these critical points.

Physically, for an observation point located away from the scatterer's surface, the critical points correspond to the points of specular reflection: there is only one such critical point on the surface of a convex scatterer. The critical points mentioned above constitute a generalization of this concept to the case in which the observation point lies *on* the scatterer's surface.

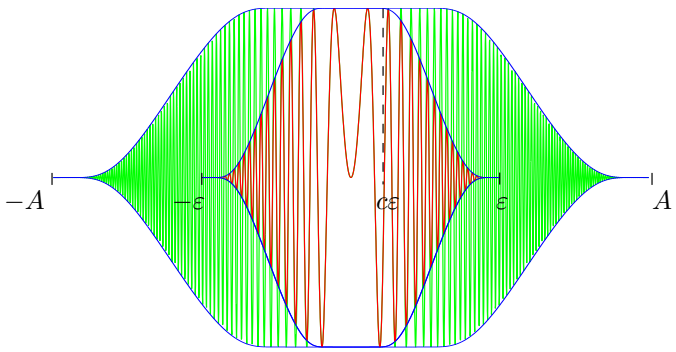


Figure A.2: Real part of functions $f_A(x)e^{ikx^p}$ and $f_\varepsilon(x)e^{ikx^p}$ with upper envelopes $f_A(x)$ and $f_\varepsilon(x)$, respectively; $p = 2$.

A.3.2 Convergent High-Frequency Integrator

To introduce our concept of localized integration let us consider the problem of integration of the one-dimensional smooth function $f_A(x)e^{ikx^p}$ depicted in Figure A.2. This discussion applies to the integrals (A.11) and (A.12) rather directly, since, via expansion of the phase ϕ in Taylor series, the oscillatory behavior of the integration kernels around their critical points is well captured by an exponential of the form e^{ikx^p} with $p = 1$ (around the kernel singularity), $p = 2$ (around the stationary points other than the shadow boundaries), or $p = 3$ (around the shadow boundary stationary points, provided the curvature does not vanish).

To state our main result concerning smooth-cutoff high-frequency integration we introduce, for real numbers $A > 0$, $0 < \varepsilon < A$ and $0 < c < 1$, explicit expressions for the functions $f_A(x)$ and $f_\varepsilon(x)$ displayed as the upper enveloping curves in Figure A.2:

$$f_A(x) = S(x, cA, A) \cdot \Big(1 - S(x, -A, -cA)\Big) \tag{A.14}$$

and

$$f_\varepsilon(x) = f_A(\frac{Ax}{\varepsilon}), \tag{A.15}$$

where

$$S(x, x_0, x_1) = \begin{cases} 1 & \text{for } x \le x_0, \\ \exp\left(\frac{2e^{-1/u}}{u-1}\right) & \text{for } x_0 < x < x_1, \quad u = \frac{x-x_0}{x_1-x_0}, \\ 0 & \text{for } x \ge x_1. \end{cases}$$

Our result now reads as follows.

Lemma A.3.1. *Let real numbers $p \geq 1$, $A > 0$, $0 < \varepsilon < A$ and $0 < c < 1$ be given, and let $f_A(x)$ and $f_\varepsilon(x)$ be defined as in equations (A.14) and (A.15) above. Then we have*

$$\int_{-A}^{A} f_A(x) e^{ikx^p}\, dx = \int_{-\varepsilon}^{\varepsilon} f_\epsilon(x) e^{ikx^p}\, dx \;+\; \mathcal{O}\Big(\big(k\varepsilon^p\big)^{-n}\Big) \quad \forall n \geq 1. \tag{A.16}$$

That is, under certain conditions on the product $k\varepsilon^p$, the integral between $-\varepsilon$ and ε of $f_\varepsilon(x)e^{ikx^p}$ is a good approximation of the integral of $f_A(x)e^{ikx^p}$ between $-A$ and A.

Proof. Defining $g_{A,\varepsilon}(x) = f_A(x) - f_\varepsilon(x)$, we obtain, for $x \geq 0$,

$$\begin{aligned} E \equiv \int_0^A f_A(x) e^{ikx^p}\, dx - \int_0^\varepsilon f_\varepsilon(x) e^{ikx^p}\, dx &= \int_{c\varepsilon}^{A} g_{A,\varepsilon}(x) e^{ikx^p}\, dx \\ &= \frac{1}{p} \int_{(c\varepsilon)^p}^{A^p} g_{A,\varepsilon}(t^{\frac{1}{p}}) t^{(\frac{1}{p}-1)} e^{ikt}\, dt. \end{aligned} \tag{A.17}$$

Integrating by parts n times and using the fact that the smooth cutoff $g_{A,\varepsilon}(x)$ vanishes together with all of its derivatives for $x = c\varepsilon$ and $x = A$, equation (A.17) becomes

$$\begin{aligned} E = \int_{(c\varepsilon)^p}^{A^p} \Big[& N_{n+1} g_{A,\varepsilon}^{(n)}(t^{\frac{1}{p}}) t^{(\frac{n+1}{p}-(n+1))} + N_n g_{A,\varepsilon}^{(n-1)}(t^{\frac{1}{p}}) t^{(\frac{n}{p}-(n+1))} \\ & + \cdots + N_2 g'_{A,\varepsilon}(t^{\frac{1}{p}}) t^{(\frac{2}{p}-(n+1))} + N_1 g_{A,\varepsilon}(t^{\frac{1}{p}}) t^{(\frac{1}{p}-(n+1))} \Big] \frac{e^{ikt}}{(ik)^n}\, dt, \end{aligned}$$

where the constants $N_{n+1}, \ldots, N_1$ depend on p, but are otherwise independent of k, ε and A. Estimating $t^{\frac{j}{p}-(n+1)} \leq D_1\big((c\varepsilon)^p\big)^{\frac{j}{p}-(n+1)}$ and $|g_{A,\varepsilon}^{(j)}| \leq D_2\varepsilon^{-j}$ on $(c\varepsilon)^p \leq t \leq A^p$, we finally obtain

$$|E| \leq D\varepsilon^{-np-p+1} k^{-n} \left| \int_{(c\varepsilon)^p}^{A^p} e^{ikt}\, dt \right|,$$

and hence (A.16). □

Error estimates for the integrals (A.11) and (A.12), similar to that of Lemma A.3.1, which can be obtained by Taylor-expanding the phase ϕ in (A.13) around the critical points, provide our criteria for the localized integration. For each target point the corresponding set of distinguished points is covered by a number of small regions, as indicated in what follows:

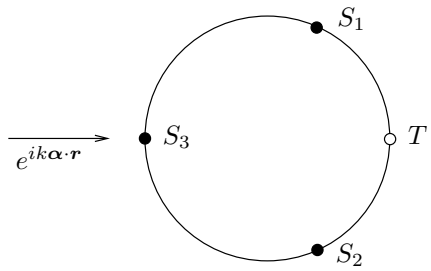


Figure A.3: Circular scatterer under plane wave incidence: target point T ($\theta_0 = 0$) and stationary phase points S_1, S_2 and S_3

Table A.1: Localized integrator, sinusoidal slow density (error on $I(\theta_0 = 0)$ using N integration points)

k	N	ε	Error
1000	2100	1.0	$1.5e{-6}$
2000	2100	0.5	$4.8e{-8}$
4000	2100	0.25	$1.2e{-7}$
8000	2100	0.125	$9.8e{-7}$
16000	2100	0.0625	$1.5e{-6}$

1. the target point is covered by a region U_t of radius proportional to the wavelength λ ($p = 1$);

2. the ℓ-th stationary point is covered by a region U_s^ℓ of radius proportional to $\sqrt{\lambda}$ ($p = 2$) or $\sqrt[3]{\lambda}$ ($p = 3$, at the shadow boundaries).

A partition of unity [14, 15] is used to smoothly split the integral over ∂D into a number of integrals over subsets of ∂D. This partition of unity is taken to be subordinated to the covering by open sets U_t and U_s^ℓ and the complement V of a closed set which is contained in and closely approximates the union of the set $U_t \cup U_s^\ell$. (In other words, the set where each of the functions making up the partition of unity is not zero is contained in one of the sets U or V.) The integral over all of ∂D is then split as a sum of integrals over V and each one of the U sets, with integrands which include the corresponding partitions of unity. The integral in the outside region V is neglected. Note that, for sufficiently small wave numbers, the U intervals cover the scatterer completely, and our high-frequency integral formulation reduces seamlessly to the original integral equation.

Let us exemplify this localized integration scheme by computing the following integral

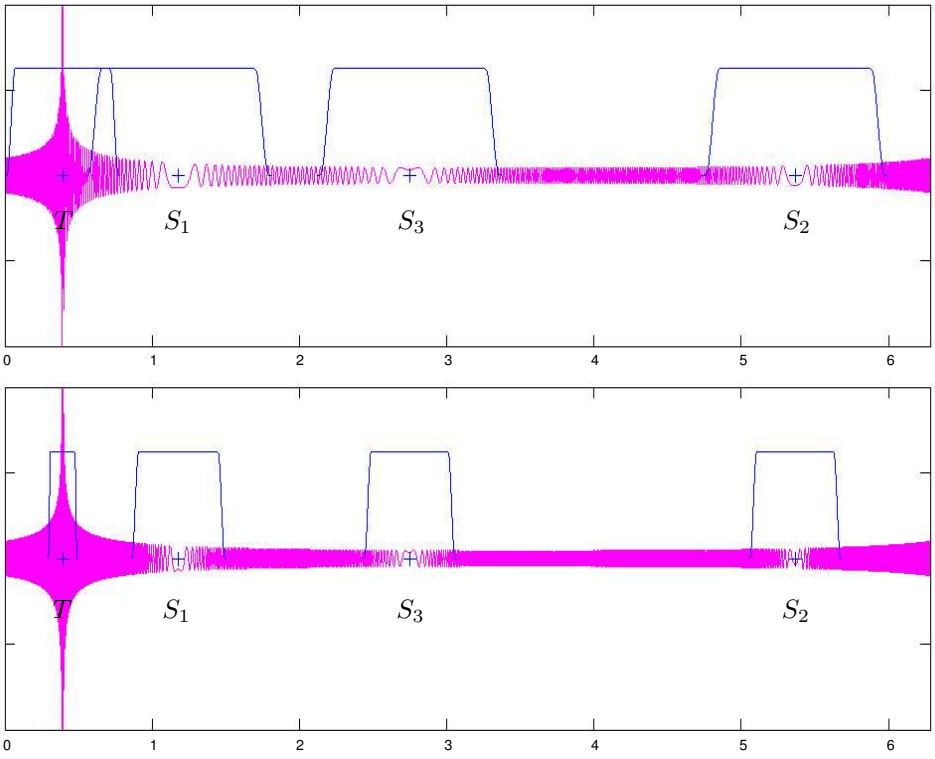


Figure A.4: Circular scatterer under plane wave incidence: smooth cutoffs around the critical points for $\theta_0 = \pi/8$, with $k = 1000$ (top) and $k = 4000$ (bottom). The quantity displayed in both graphs is the real part of the integrand in (A.18), divided by $\cos(\theta)$.

on a circle of unit radius, centered at the origin (see Figure A.3):

$$I(\theta_0) = \int_0^{2\pi} \left[H_0^1(k|\boldsymbol{r}(\theta_0) - \boldsymbol{r}(\theta)|) e^{ik\boldsymbol{\alpha}\cdot(\boldsymbol{r}(\theta)-\boldsymbol{r}(\theta_0))} \right] \cos(\theta)\, d\theta, \tag{A.18}$$

with $\boldsymbol{r}(\theta) = (\cos\theta, \sin\theta)$. Equation (A.18) corresponds to the two-dimensional single layer potential in the integral equation (A.7), with the unknown density substituted by $\cos(\theta)$. Table A.1 demonstrates the fixed accuracy of the integrator for $\theta_0 = 0$ and $\alpha = (1, 0)$ throughout the frequency spectrum, using a *fixed* number of integration points for all values of k. Figure A.4 illustrates the covering of the critical points by the intervals U_t and U_s^ℓ ($\ell = 1, 2, 3$) for the target point $\theta_0 = \pi/8$, and highlights the variation of the size of the local integration intervals as the frequency increases.

Table A.2: Interpolation of the slow density (number of coefficients in the Fourier expansion of $\mu_{\text{slow}}^{\text{exact}}$ for a circular scatterer, with and without the change of variables depicted in Figure A.5)

ka	$\mu_{\text{slow}}^{\text{exact}}(\eta_k(\theta))$	$\mu_{\text{slow}}^{\text{exact}}(\theta)$
100	110	110
1000	220	230
10000	280	310
100000	280	350
1000000	280	> 500

A.3.3 Shadow Boundaries

Shadow boundaries (where $\boldsymbol{\alpha} \cdot \boldsymbol{\nu} = 0$, see Figure A.5) require special consideration. Indeed, in order to represent μ_{slow} within a fixed error tolerance by means of a frequency-independent discretization density, a cubic root singularity inherent in the slow density around such boundaries needs to be accounted for appropriately. Figure A.5 (lower left) illustrates the $k^{1/3}$ increase of the slopes of the slow density phases $\phi = \phi_k(\theta)$ around the shadow boundary as k increases. Figure A.5 (lower right), in turn, displays the effect of the change of variables $\eta_k = \eta_k(\theta)$, displayed in Figure A.5 (upper right), that we use to compensate for this effect. Table A.2 compares the number of Fourier modes required to represent the closed form, exact slow density $\mu_{\text{slow}}^{\text{exact}}$ for a circular scatterer, within a certain error tolerance, with and without the introduction of this change of variables around the shadow boundary. We see that, after the change of variables has been applied, $\mu_{\text{slow}}^{\text{exact}}$ can be represented, with a fixed accuracy, by Fourier series with a fixed number of terms for arbitrarily large wave numbers k.

A.4 Spectral Implementation

A.4.1 High-Order Interpolation

In view of the discussion of Section A.2.1 and Section A.3.3, μ_{slow} in (A.10) can be obtained, within a prescribed error tolerance, through interpolation from a fixed (independent of frequency) number of discretization points.

In our implementation, these points are associated with the nodes of Cartesian grids discretizing one or more (overlapping) patches covering the scatterer surface, as proposed

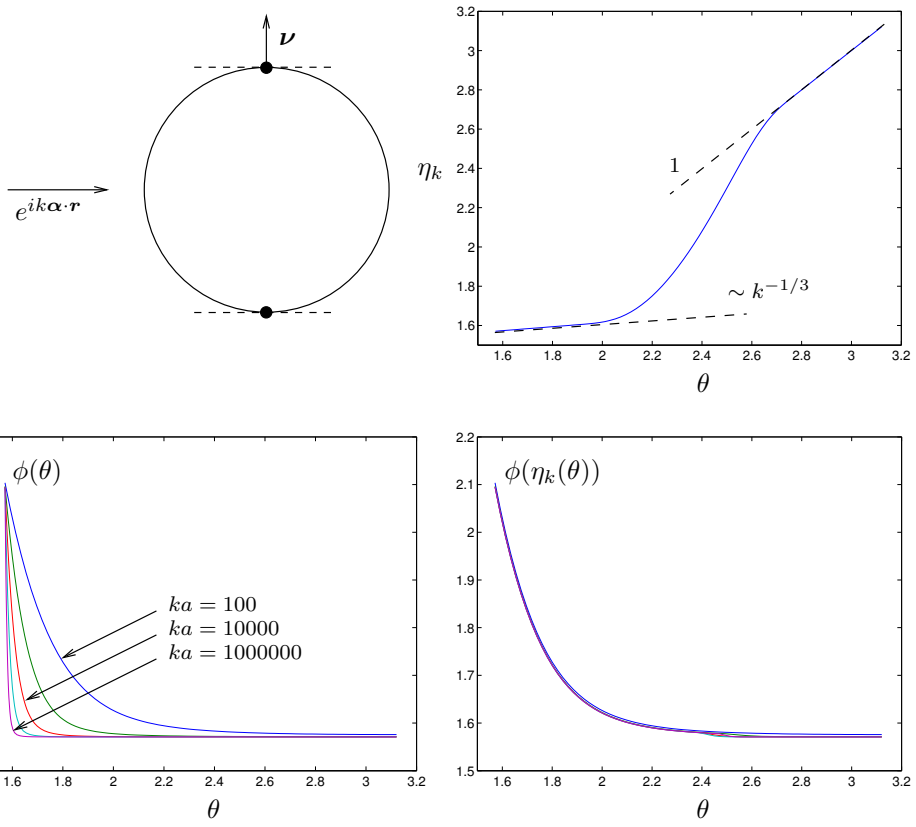


Figure A.5: Top left: shadow boundary for the circular cylinder. Top right: cubic root change of variables. Bottom left and bottom right: variation of the phase of $\mu_{\text{slow}}^{\text{exact}}$ at the shadowing point before and after application of the change of variables, respectively, for $k = 100$ to $k = 1000000$.

by [14, 15]. Fast interpolations of very high order can then be obtained using refined FFTs and polynomial off-grid interpolation [8]:

1. Using one- or two-dimensional FFTs (in two and three dimensions, respectively), construct a Fourier series for each interpolation patch. Thanks to the partition of unity subordinated to these patches, the densities are smooth and vanish on the patch boundaries and the convergence of these Fourier series is high-order;

2. Use FFTs to evaluate the Fourier series (and possibly their derivatives) on much refined, but still equispaced, grids. The actual choice of the refinement factor is based on a trade-off between computational times and accuracy;

3. Use the density values on the refined grids to construct one or more local interpolation polynomials per original grid interval.

In our numerical examples we used a 32-fold refinement of the original grids and cubic splines for the local polynomial interpolation. Clearly the convergence of this algorithm is only fourth order in the sub-grid spacing, but the error it introduces (compared to an explicit evaluation of the Fourier series) is several orders of magnitude smaller than the overall error on the solution of the problems we considered (see Section A.5). If true super-algebraic convergence is required one could replace the cubic splines by Chebyshev interpolation, or even, at the expense of significantly slower numerics, by an unequally spaced FFT algorithm [22].

Note that, for practical problems (where the geometrical description of the scatterer is not known analytically), a high-order surface representation (such as that described in [9]) is also required to preserve the high-order convergence of the method.

A.4.2 Trapezoidal-Rule Integration

The integral in the region U_t (see Section A.3.1), which contains the kernel singularity, is evaluated by means of a discretization with a mesh-size proportional to λ. Our choice of singular integrator is that described by [20] in the two-dimensional case and by [14, 15] in the three-dimensional case. These methods provide high-order quadrature for the singular integrands arising in the integral equations under consideration.

The integral in the region U_s^ℓ, in turn, is evaluated by means of the trapezoidal rule with a discretization mesh-size proportional to $\sqrt{\lambda}$ or $\sqrt[3]{\lambda}$.

Table A.3: Scattering of an incident plane wave on a circular cylinder of radius a

25 unknowns, $\varepsilon = \varepsilon_{\text{ref}}$			
ka	GMRES iterations	Error	CPU time
1	9	$1.0e{-}12$	$< 1s$
10	11	$1.6e{-}4$	$< 1s$
100	13	$9.3e{-}4$	$3s$
1000	13	$8.3e{-}3$	$5s$
10000	15	$1.0e{-}2$	$6s$
100000	14	$1.1e{-}2$	$6s$
100 unknowns, $\varepsilon = 5\varepsilon_{\text{ref}}$			
ka	GMRES iterations	Error	CPU time
1	9	$1.0e{-}12$	$< 1s$
10	17	$3.0e{-}11$	$5s$
100	22	$1.5e{-}5$	$11s$
1000	25	$3.1e{-}5$	$2m30s$
10000	27	$8.4e{-}5$	$3m12s$
100000	30	$8.8e{-}5$	$3m43s$

In all cases, the values of the slow densities at the integration points are obtained through interpolation from the fixed discretization mesh mentioned above. Note that, because of the smooth cutoffs used, all integrands are smooth periodic functions—for which the trapezoidal rule gives rise to high-order convergence. Also note that a special procedure is necessary to guarantee that the non-empty intersections occurring between the various U sets defined above do not cause difficulties: if the sets have identical discretizations, they are simply merged and the corresponding elements of the partition of unity are summed; otherwise, in a recursive manner, the integral on the set having the finer discretization is computed completely, and its partition of unity subtracted from the other sets.

A.5 Numerical Results

A matrix-free iterative solver has been implemented by utilizing the two-dimensional version of the high-frequency integrator described in the preceding sections in conjunction with the GMRES algorithm [53]. Table A.3 shows results produced by means of this two-dimensional solver on a $1.5\,GHz$ PC, applied to a circular cylinder of radius a. Errors were computed

by comparison with an exact solution for the integral equation, and defined as

$$\left\{\int_{\partial D} |\mu_{\text{slow}}^{\text{exact}}(\boldsymbol{r}) - \mu_{\text{slow}}(\boldsymbol{r})|^2 \, ds(\boldsymbol{r})\right\}^{1/2} \Big/ \left\{\int_{\partial D} |\mu_{\text{slow}}^{\text{exact}}(\boldsymbol{r})|^2 \, ds(\boldsymbol{r})\right\}^{1/2}. \tag{A.19}$$

This example illustrates the asymptotically bounded complexity of our approach: the error for $k = 1000$ is almost identical to that for $k = 100000$, using the same number of unknowns and the same number of integration points. The high-frequency solver is well conditioned and requires a small number of GMRES iterations for arbitrarily large wave numbers, leading to nearly identical computation times for all values of $k > 1000$.

The results given in the upper half of Table A.3 were obtained using 25 discretization points for the slow density μ_{slow} and a local integration interval size $\varepsilon_{\text{ref}} = 600(ka)^{-1}$, with 800 integration points per interval. As can be seen in the table, these parameters led to a maximum error of about 1% throughout the frequency spectrum, in total computational times of 6 seconds.

In order to obtain two more digits of accuracy, the number of discretization points was increased by a factor of 4 and the size of the integration intervals multiplied by 5 (see bottom half of Table A.3). In this case, using 4200 integration points per interval led to an error smaller that $1.e{-}4$, in less than 4 minutes CPU time.

Thanks to the very accurate interpolation scheme presented in Section A.4.1, the overall error of the algorithm is dominated by two independent sources: on the one hand, the discretization error of μ_{slow} and, on the other hand, the error introduced by the local integrator. To obtain a given error tolerance *with the smallest computational cost*, a compromise has to be found between the number of unknowns and the size of the local integration intervals. For example, with $\varepsilon = \varepsilon_{\text{ref}}$ numerical tests showed that increasing the number of unknowns beyond 25 does not improve the accuracy of the solution beyond the values shown in Table A.3, whereas, with $\varepsilon = 5\varepsilon_{\text{ref}}$, using only 25 instead of 100 unknowns would result in a global error of $7e{-}3$ (instead of $8.8e{-}5$).

As a comparison, the computational time required by the direct algorithm described in [20, p. 66] to yield an error of $4.8e{-}3$ for $ka = 1000$ is 36 minutes—using the same linear algebra solver as our high-frequency solver and the same computer; our solver produces a comparable accuracy in a 5 second calculation. Since the computational time for the direct solver grows quadratically with the wavenumber, that method would require 250 days to

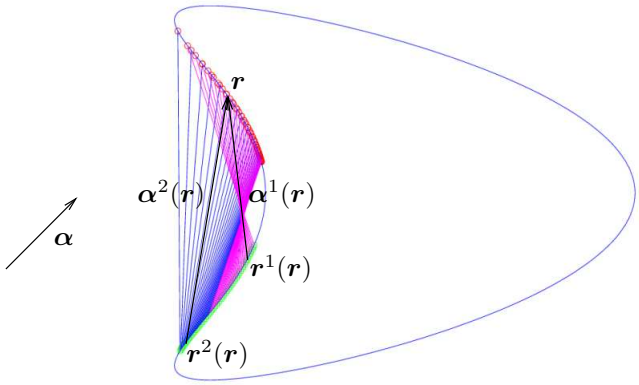


Figure A.6: A multiple scattering configuration: the kite-shaped scatterer of equation (A.21)

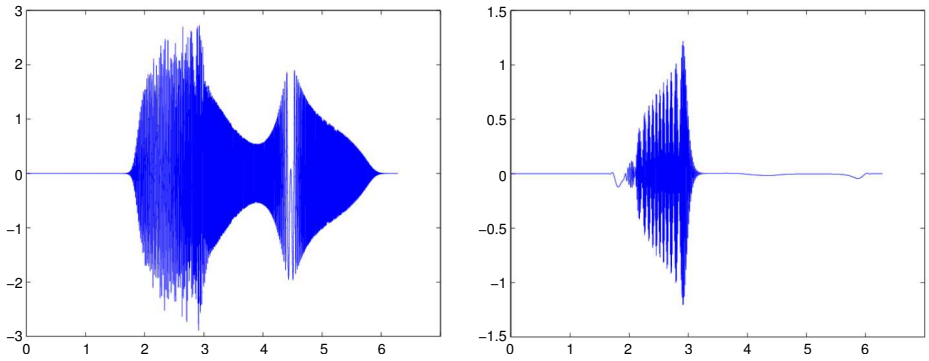


Figure A.7: Real parts of $\mu(\boldsymbol{r}(\theta))$ and $\mu(\boldsymbol{r}(\theta))/e^{ik\boldsymbol{\alpha}\cdot\boldsymbol{r}(\theta)}$ on the kite-shaped scatterer, for $k = 800$

produce a solution for $ka = 100000$ with the same error. The corresponding solution presented in the upper portion of Table A.3 was obtained in a mere 6 second calculation.

A.6 Extension to Non-Convex Scatterers

For a non-convex obstacle the ansatz (A.5) may not be valid; indeed, this ansatz is not valid if the incident field is such that its ray theory approximation gives rise to multiple reflections, as shown in Figure A.6. This fact can be understood easily: in the case of Figure A.6, for example, the reflections in the lower part of the concave region act as sources of an "incident" field for the upper portion of the concavity which should be accounted for in the ansatz for the density μ.

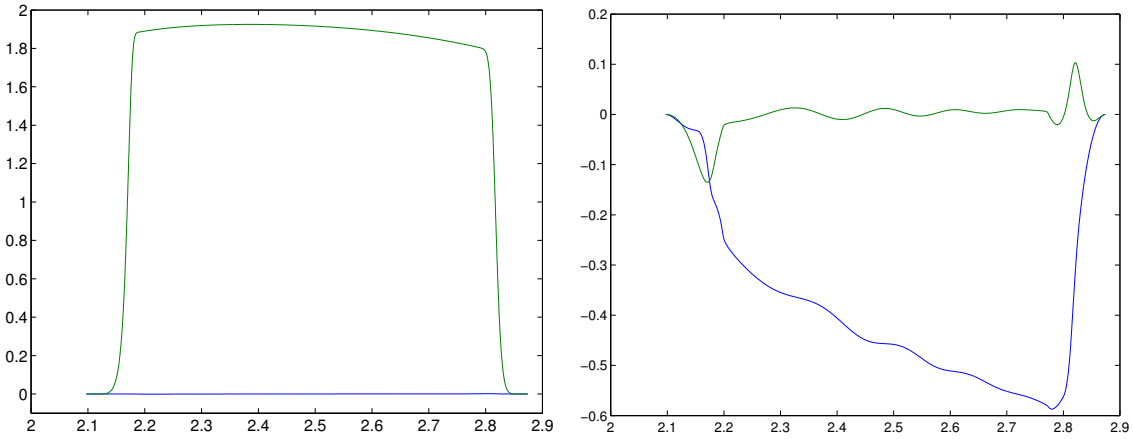


Figure A.8: Real and imaginary parts of the slow densities $\mu^0_{\text{slow}}(\boldsymbol{r}(\theta))$ and $\mu^1_{\text{slow}}(\boldsymbol{r}(\theta))$ on the kite-shaped scatterer, for $k = 800$

A correct version of the ansatz for the scatterer of Figure A.6 is

$$\begin{aligned}\mu(\boldsymbol{r}) = \mu^0_{\text{slow}}(\boldsymbol{r})e^{ik\boldsymbol{\alpha}\cdot\boldsymbol{r}} + \mu^1_{\text{slow}}(\boldsymbol{r})e^{ik\boldsymbol{\alpha}^1(\boldsymbol{r})\cdot(\boldsymbol{r}-\boldsymbol{r}^1(\boldsymbol{r}))}\, e^{ik\boldsymbol{\alpha}\cdot\boldsymbol{r}^1(\boldsymbol{r})} + \\ \mu^2_{\text{slow}}(\boldsymbol{r})e^{ik\boldsymbol{\alpha}^2(\boldsymbol{r})\cdot(\boldsymbol{r}-\boldsymbol{r}^2(\boldsymbol{r}))}\, e^{ik\boldsymbol{\alpha}\cdot\boldsymbol{r}^2(\boldsymbol{r})}\end{aligned} \tag{A.20}$$

where $\mu^1_{\text{slow}}(\boldsymbol{r})$ and $\mu^2_{\text{slow}}(\boldsymbol{r})$ are compactly supported in the upper part of the concavity and where $\boldsymbol{\alpha}^1(\boldsymbol{r})$ and $\boldsymbol{\alpha}^2(\boldsymbol{r})$ are unit vectors of direction given by the geometrical optics rays. Note that, indeed, two slow densities arise as a result of the reflections from the lower portions of the concavity, since points in the upper part of the concavity are illuminated by two reflections from the lower part. The algorithm for the non-convex case can be completed by iteratively applying the algorithm for convex scatterers to this modified type of ansatz—the details of such an implementation will be presented elsewhere.

Figure A.7 displays the solution, for $k = 800$, of the multiple scattering by the kite-shaped scatterer of Figure A.6, described by the parametric representation

$$\boldsymbol{r}(\theta) = (\cos\theta + 0.65\cos 2\theta - 0.65, 1.5\sin\theta), \quad 0 \le \theta \le 2\pi. \tag{A.21}$$

The actual solution $\mu(\boldsymbol{r})$ is shown on the left, while $\mu(\boldsymbol{r})/e^{ik\boldsymbol{\alpha}\cdot\boldsymbol{r}}$ (i.e., the solution that would be obtained using the simple ansatz (A.5)) is shown on the right and clearly demonstrates the inapplicability of (A.5) in the multiple scattering case. Figure A.8, in turn, demonstrates that slow densities $\mu^0_{\text{slow}}(\boldsymbol{r})$ and $\mu^1_{\text{slow}}(\boldsymbol{r})$ can indeed be obtained in such a way that the

ansatz (A.20) holds in the multiple scattering case.

A.7 Conclusions

We have presented a high-frequency integrator which can evaluate all the relevant integrals within a given error tolerance through a fixed number of operations—independently of frequency. When used in conjunction with a Krylov subspace method such as GMRES, numerical experiments have shown a good conditioning of the discrete operator, requiring a virtually constant number of GMRES iterations to attain a fixed error prescription as the frequency increases. The overall solver has subsequently been shown to be able to solve scattering problems within a prescribed error tolerance for arbitrarily small wavelengths, within a fixed computing time.

A.8 Evaluation of Stationary Points

Note: this section originally appeared as an appendix in [12].

We consider two-dimensional obstacles whose boundaries admit polar parameterizations

$$r = r(\theta), \quad 0 \leq \theta \leq 2\pi.$$

Let the phase of the incident wave be given by

$$ik\boldsymbol{\alpha} = ik(\cos(\varphi), \sin(\varphi))$$

and consider the total phase

$$ik\phi = ik\left[|\boldsymbol{y} - \boldsymbol{x}| + \boldsymbol{\alpha} \cdot \boldsymbol{y}\right]$$

where $\boldsymbol{x}$ and $\boldsymbol{y}$ are arbitrary points on the boundary of the obstacle, the *target* and *source* points respectively. (The total phase is obtained as the phase in the $k \to \infty$ asymptotic expression of $G(\boldsymbol{x}, \boldsymbol{y})e^{ik\boldsymbol{\alpha}\cdot\boldsymbol{y}}$). Without loss of generality we may assume $\varphi = 0$. Using polar coordinates

$$\boldsymbol{x} = r(\theta_0)e^{i\theta_0} \quad \text{and} \quad \boldsymbol{y} = r(\theta)e^{i\theta},$$

we have

$$\phi = \phi(\theta) = \rho + r(\theta)\cos(\theta) = \sqrt{r(\theta_0)^2 + r(\theta)^2 - 2r(\theta_0)r(\theta)\cos(\theta - \theta_0)} + r(\theta)\cos(\theta).$$

The stationary points then correspond to the solutions of

$$0 = \phi'(\theta) = \frac{r(\theta)\dfrac{dr}{d\theta}(\theta) - r(\theta_0)\dfrac{dr}{d\theta}(\theta)\cos(\theta - \theta_0) + r(\theta_0)r(\theta)\sin(\theta - \theta_0)}{\sqrt{r(\theta_0)^2 + r(\theta)^2 - 2r(\theta_0)r(\theta)\cos(\theta - \theta_0)}} + \frac{dr}{d\theta}(\theta)\cos(\theta) - r(\theta)\sin(\theta)$$

in the interval $[0, 2\pi)$. The solution of this nonlinear equation can be obtained in $\mathcal{O}(N)$ operations by means of Newton's method. For the particular case of the circle (i.e., $r(\theta) \equiv R$) the condition for stationarity is

$$0 = \phi'(\theta) = \frac{R^2 \sin(\theta - \theta_0)}{\sqrt{2R^2(1 - \cos(\theta - \theta_0))}} - R\sin(\theta)$$

and a closed form can be given:

$$0 \le \theta - \theta_0 = \pi - 2\theta_0 + 4n\pi \le 2\pi \quad \text{or} \quad 0 \le \theta - \theta_0 = \frac{(\pi - 2\theta_0)}{3} + \frac{4}{3}\pi n \le 2\pi.$$

Appendix B

Code Verification

To verify our code, we apply it and the solver from [13] to three scattering cases. Table B.1 lists the physical variables for the three configurations, all of which have gratings of the form $f(x) = \frac{h}{2}\cos(2\pi x)$. Cases 1 and 2 both have only simple reflections (see Figure 2.10), but Case 1 has TE/sound-soft scattering while Case 2 has TM/sound-hard scattering; Case 3 has both multiple reflections and shadowing with TE/sound-soft scattering. Thus, for Cases 1 and 2 we compute $\mu_2(x)$ from the appropriate scattering equations, while for Case 3 we compute $\mu_1(x)$. In all three problems we use $A_{\text{sp}} = \frac{7}{8} = 0.875$, while the numerical parameters n_{t}, n_{i} and A are varied and are listed in Table B.1.

For these cases our code demonstrates itself accurate to machine precision. Not only are the energy balance errors indicative of this accuracy (see Table B.1), but the computed values for each of the scattering efficiencies e_n agree between the two codes (see Tables B.2 and B.3 for Cases 1 and 2 and Figure B.1 for Case 3).

We also compute the scattering efficiencies of three of the systems in [17], the work which describes a high-frequency method that we briefly reviewed in Section 2.3.2.1. We denote these systems as Cases 4, 5 and 6; Cases 4 and 5 have a grating of the form $f(x) = \frac{h}{2}\cos(2\pi x)$ and have TE/sound-soft and TM/sound-hard scattering, respectively, while Case 6 has a

case	h	$\frac{k}{2\pi}$	θ	n_{t}	n_{i}	A	e.b. error (this work)	e.b. error ([13] code)
1	0.025	1.5	0°	12	12×8	350	4.4×10^{-16}	1.1×10^{-15}
2	0.025	1.5	30°	12	12×8	700	1.1×10^{-16}	4.4×10^{-16}
3	0.25	10.0	75°	96	96×3	750	1.7×10^{-15}	8.9×10^{-16}

Table B.1: Physical quantities, numerical parameters and results for the cases computed by our method and the method of [13]. The energy balance errors are listed as "e.b. error."

n	e_n	differences
-1	$1.026215905707786 \times 10^{-2}$	1.3×10^{-16}
0	$9.794756818858454 \times 10^{-1}$	-1.8×10^{-15}
1	$1.026215905707786 \times 10^{-2}$	1.1×10^{-16}

Table B.2: Case 1 efficiencies and the differences in their computed values between this work and the solver of [13]

n	e_n	differences
-2	$8.930278583943842 \times 10^{-5}$	-2.3×10^{-16}
-1	$1.882452296791681 \times 10^{-2}$	3.3×10^{-16}
0	$9.810861742462433 \times 10^{-1}$	2.2×10^{-16}

Table B.3: Case 2 efficiencies and the differences in their computed values between this work and the solver of [13]

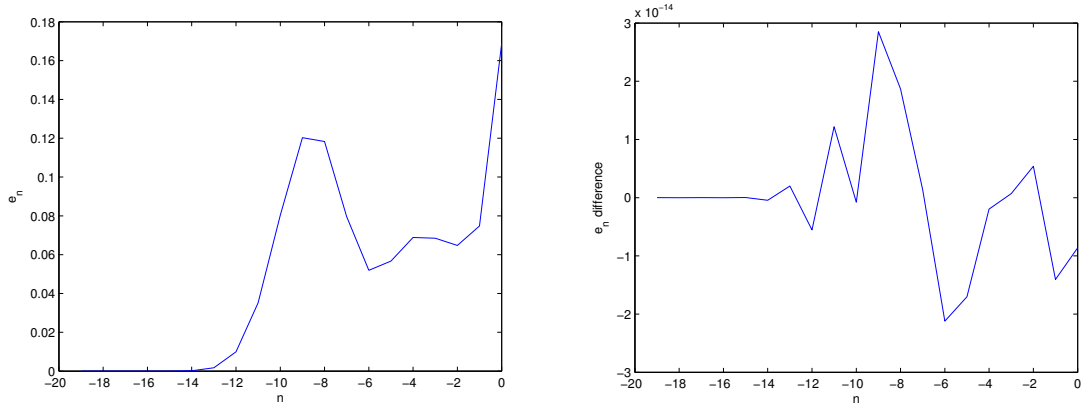


Figure B.1: Case 3 efficiencies and the differences in their computed values between this work and the solver of [13]

case	h	λ	θ	n_{t}	n_{i}	A	energy balance error
4	0.025	0.025	30°	24	24×8	20	8.9×10^{-16}
5	0.025	0.025	30°	24	24×8	20	1.1×10^{-15}
6	0.02	0.04	0°	56	56×4	10	2.2×10^{-16}

Table B.4: Physical quantities, numerical parameters and results for the cases from [17]

grating of the form

$$f(x) = \frac{h}{2}\left[-\cos(2\pi x) + 0.35\cos(4\pi x) - 0.035\cos(6\pi x)\right]$$

and TE/sound-soft scattering. Table B.4 indicates the physical parameters for these cases as well as the numerical parameters (in all cases $A_{\mathrm{sp}} = 0.875$) and the energy balance errors of our code (computing $\mu_2(x)$ in each case); Figure B.2 has plots of the scattering efficiencies for these cases. Tables B.5–B.7 list additional information about particular scattering efficiencies. They include reference values and relative errors of the highest-order computed efficiencies as listed in [17] along with the relative errors of the computed efficiencies of our code. Again, our code demonstrates itself accurate to machine precision.

We note that for these three cases from [17] the values of k are Wood Anomaly wavenumbers (Remark 2.1.5). For Cases 4 and 5, $k = 40 \times 2\pi$ and $\theta = 30°$, so

$$\alpha_n = k\sin(\theta) + n\frac{2\pi}{L} = 20.0 \times 2\pi + n \times 2\pi \tag{B.1}$$

implies that

$$\beta_{20} = \sqrt{k^2 - \alpha_{20}^2} = 0 \tag{B.2}$$

and

$$\beta_{-60} = \sqrt{k^2 - \alpha_{-60}^2} = 0. \tag{B.3}$$

For Case 6, $k = 25 \times 2\pi$ and $\theta = 0°$, so $\beta_{-25} = \beta_{25} = 0$. Thus, our method can be applied to these configurations even though the periodic Green's function is not well defined for these wavenumbers (Section 2.1.3.1).

Remark B.0.1. *The computational times of our code were under* 2 *seconds for Cases* 1 *and* 2*, about* 44 *seconds for Case* 3 *and under* 0.5 *seconds for Cases* 4–6.

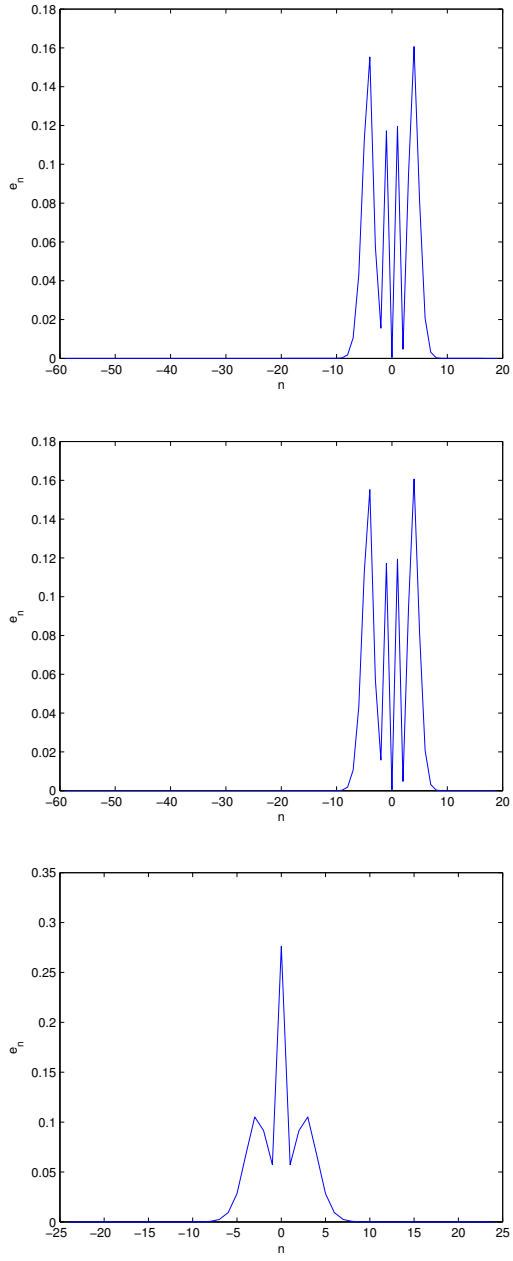


Figure B.2: Case 4 (top), Case 5 (middle) and Case 6 (bottom) efficiencies

n	e_n (reference)	relative error (method of [17])	relative error (this work)
0	$7.538669511479800 \times 10^{-4}$	2.4×10^{-13}	7.2×10^{-14}
1	$1.194293110668300 \times 10^{-1}$	2.2×10^{-14}	4.2×10^{-15}
2	$4.713900020760300 \times 10^{-3}$	3.3×10^{-14}	2.6×10^{-14}
3	$9.472951023686101 \times 10^{-2}$	4.0×10^{-15}	2.2×10^{-14}
4	$1.606247510782500 \times 10^{-1}$	8.6×10^{-15}	1.2×10^{-14}
5	$8.121747375826800 \times 10^{-2}$	7.9×10^{-15}	3.5×10^{-14}
6	$2.068175899532900 \times 10^{-2}$	4.4×10^{-15}	2.1×10^{-14}
7	$3.171379802403400 \times 10^{-3}$	5.3×10^{-15}	5.3×10^{-15}

Table B.5: Efficiencies and relative errors of Case 4

n	e_n (reference)	relative error (method of [17])	relative error (this work)
0	$6.978718873398379 \times 10^{-4}$	1.6×10^{-15}	1.0×10^{-13}
1	$1.193803726254851 \times 10^{-1}$	9.3×10^{-16}	8.5×10^{-15}
2	$4.854671479355886 \times 10^{-3}$	5.4×10^{-16}	1.2×10^{-14}
3	$9.427330239288337 \times 10^{-2}$	2.9×10^{-16}	4.7×10^{-15}
4	$1.606619051666006 \times 10^{-1}$	5.2×10^{-16}	6.9×10^{-16}
5	$8.146471443830940 \times 10^{-2}$	0.0×10^{-16}	6.8×10^{-16}
6	$2.079411505463193 \times 10^{-2}$	1.0×10^{-15}	1.0×10^{-15}
7	$3.195973191313253 \times 10^{-3}$	1.9×10^{-15}	5.4×10^{-14}

Table B.6: Efficiencies and relative errors of Case 5

n	e_n (reference)	relative error (method of [17])	relative error (this work)
0	$2.762105662320035 \times 10^{-1}$	2.4×10^{-15}	4.0×10^{-16}
1	$5.735818584364873 \times 10^{-2}$	6.0×10^{-16}	3.9×10^{-15}
2	$9.154897389472935 \times 10^{-2}$	6.7×10^{-15}	6.4×10^{-15}
3	$1.051875097051952 \times 10^{-1}$	9.2×10^{-16}	9.2×10^{-16}
4	$6.713521833646909 \times 10^{-2}$	2.1×10^{-16}	8.7×10^{-15}
5	$2.830374622545111 \times 10^{-2}$	6.7×10^{-15}	2.5×10^{-15}
6	$9.270117932865375 \times 10^{-3}$	3.0×10^{-15}	4.7×10^{-14}
7	$2.435385416440963 \times 10^{-3}$	1.8×10^{-16}	7.1×10^{-16}

Table B.7: Efficiencies and relative errors of Case 6

Appendix C

Additional TM/Sound-Hard Results

We re-examine six of the examples discussed in Chapter 4—keeping the scattering surfaces as well as the incident fields' wavenumbers and incidence angles of these cases the same but changing the type of scattering from TE/sound-soft to TM/sound-hard. The purpose of this study is to demonstrate that the computational accuracy of our solver for these TM/sound-hard problems is virtually the same as it is for the corresponding TE/sound-soft problems.

The configurations selected all contain grating profiles of the form $f(x) = \frac{h}{2}\cos(2\pi x)$, except for the one (here denoted as Case 2) which has the "multi-scale" surface $f(x) = \frac{h}{2}\left[\cos(2\pi x) + 0.04\sin(50\pi x)\right]$. A wide variety of heights h, wavenumbers k and incidence angles θ are included in this study; Table C.1 lists these physical quantities as well as the sections from Chapter 4 in which the original TE/sound-soft cases are discussed. In applying our solver to these TM/sound-soft cases, we use representations of the "densities" (either $\mu_1(x)$ or $\mu_2(x)$) and sets of numerical parameters (n_{t}, n_{i}, A_{sp} and A) that were used to generate the TE/sound-soft results. The representations are given in Table C.1, and the sets of numerical parameters employed are detailed in Table C.2. Note that we use two sets (denoted as "(a)" and "(b)") of numerical parameters for both Case 1 and Case 2 in order to make comparisons at distinct levels of accuracy.

Table C.3 describes the energy balance errors achieved by our solver (the TE/sound-soft data are taken from the relevant sections of Chapter 4). Indeed, given a particular grating profile, wavenumber, incidence angle, representation of the "density" and set of numerical parameters, the method of this thesis computes the resulting scattering efficiencies at around

case	section	rep.	h	$\frac{k}{2\pi}$	θ
1	4.1.1.2	$\mu_1(x)$	0.25	10	75°
2	4.2.2.2	$\mu_2(x)$	0.025	10.5	30°
3	4.3.1	$\mu_2(x)$	0.025	10000	10°
4	4.3.2	$\mu_1(x)$	4.0	10	10°
5	4.3.3	$\mu_1(x)$	0.025	10	87°
6	4.3.4.1	$\mu_2(x)$	0.025	100	30°

Table C.1: Physical quantities and choices of representation for the densities of the TM cases, plus the sections in Chapter 4 discussing the corresponding TE cases

case	n_t	n_i	A_sp	A
1(a)	32	32×2	0.875	750
1(b)	96	96×3	0.875	750
2(a)	64	64×1	0.875	2
2(b)	220	220×1	0.875	200
3	16	16×2000	0.0175	0.02
4	360	360×2	0.875	800
5	24	24×6	0.875	11000
6	432	432×1	0.875	4

Table C.2: Numerical parameters for the TM cases

the same level of accuracy whether TE/sound-soft scattering or TM/sound-hard scattering is occurring.

Remark C.0.2. *Cases* 4 *and* 5 *of Appendix B together form another example of our solver producing similar energy balance errors for the two types of scattering if all else is equal.*

case	energy balance error (TE)	energy balance error (TM)
1(a)	1.0×10^{-9}	6.7×10^{-9}
1(b)	1.7×10^{-15}	1.9×10^{-14}
2(a)	1.1×10^{-4}	1.6×10^{-4}
2(b)	0.0×10^{-16}	3.3×10^{-16}
3	7.8×10^{-16}	7.5×10^{-15}
4	4.6×10^{-13}	4.7×10^{-12}
5	3.3×10^{-16}	2.4×10^{-13}
6	2.9×10^{-15}	3.6×10^{-15}

Table C.3: TE and TM energy balance errors

Appendix D

Rounding Errors

Due to the finite precision of the computational data (16 digits for double precision data), rounding errors are inherent to any numerical solver. For our method, they accumulate to some extent as we continue to increase the integration window size A. Taking the simple-reflection case from Section 4.1.2 (using the same reference solution), we demonstrate in Table D.1 that, with fixed discretization values $n_\mathrm{t} = n_\mathrm{i} = 48$, the energy balance error and the maximum absolute error increase as A is increased far above 30. As shown in Figure D.1, the cancellation errors for the $A = 30000$ solution generate "noise" which artificially boosts the size of the insignificant Fourier coefficients of the $A = 30$ solution, and this affects the accuracy of the computed efficiencies. These errors are of minor importance, however.

A	energy balance error	max. abs. error
1	6.6×10^{-5}	7.5×10^{-5}
2	2.8×10^{-6}	1.2×10^{-6}
5	3.0×10^{-8}	1.1×10^{-8}
10	9.6×10^{-11}	2.8×10^{-11}
20	2.8×10^{-14}	2.3×10^{-14}
30	8.9×10^{-16}	2.4×10^{-16}
300	6.4×10^{-15}	5.2×10^{-15}
3000	2.9×10^{-14}	7.7×10^{-14}
30000	1.3×10^{-13}	8.5×10^{-13}

Table D.1: Results for various A ($n_\mathrm{t} = n_\mathrm{i} = 48$) for the simple-reflection case of Section 4.1.2

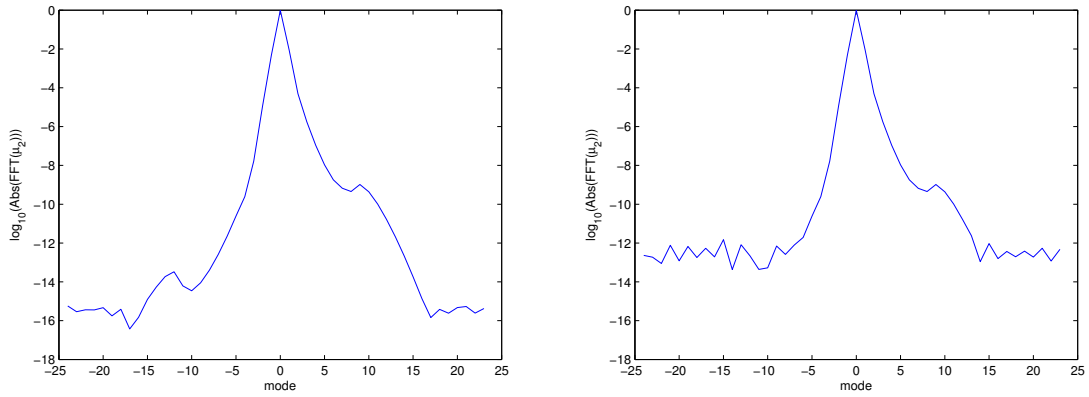


Figure D.1: Fourier amplitudes for the $A = 30$ solution (left) and the $A = 30000$ solution (right) for the simple-reflection case of Section 4.1.2

Bibliography

[1] K.R. Aberegg and A.F. Peterson, Application of the integral equation-asymptotic phase method to two-dimensional scattering, *IEEE Transactions on Antennas and Propagation*, 43(5), 534–537, 1995.

[2] T. Abboud, J.-C. Nédélec and B. Zhou, Méthode des équations intégrales pour les hautes fréquences, *Comptes rendus de l'Académie des sciences, Série 1, Mathématique*, 318(2), 165–170, 1994.

[3] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, New York, Dover, 1965.

[4] T. Arens, S.N. Chandler-Wilde and J.A. DeSanto, On Integral Equation and Least Squares Methods for Scattering by Diffraction Gratings, *Communications in Computational Physics*, 1(6), 1010–1042, 2006.

[5] N. Bleistein and R.A. Handelsman, *Asymptotic Expansions of Integrals*, New York, Dover, 1986.

[6] P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces*, New York, Macmillan, 1963.

[7] D. Bouche, F. Molinet and R. Mittra, *Asymptotic Methods in Electromagnetics*, Berlin, Springer-Verlag, 1997.

[8] J.P. Boyd, A fast algorithm for Chebyshev, Fourier & sinc interpolation onto an irregular grid, *Journal of Computational Physics*, 103(22), 243–257, 1992.

[9] O.P. Bruno, Fast, high-order, high-frequency integral methods for computational acoustics and electromagnetics, in *Topics in computational wave propagation: direct and inverse problems, Lecture Notes in Computational Science and Engineering, No. 31*, 43–82, Berlin, Springer-Verlag, 2003.

[10] O.P. Bruno, Wave scattering by inhomogeneous media: efficient algorithms and applications, *Physica B: Condensed Matter*, 338(11), 67–73, 2003.

[11] O.P. Bruno and C.A. Geuzaine, An $\mathcal{O}(1)$ integration scheme for three-dimensional surface scattering problems, *Journal of Computational and Applied Mathematics*, 204(2), 463–476, 2007.

[12] O.P. Bruno, C.A. Geuzaine, J.A. Monro Jr. and F. Reitich, Prescribed error tolerances within fixed computational times for scattering problems of arbitrarily high frequency: the convex case, *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, 362(1816), 629–645, 2004.

[13] O.P. Bruno and M.C. Haslam, High-order solution of the scattering problem for one-dimensional perfectly-conducting periodic surfaces, (*in process*).

[14] O.P. Bruno and L.A. Kunyansky, A fast, high-order algorithm for the solution of surface scattering problems: basic implementation, tests, and applications, *Journal of Computational Physics*, 169, 80–110, 2001.

[15] O.P. Bruno and L.A. Kunyansky, Surface scattering in three dimensions: an accelerated high-order solver, *Proceedings of the Royal Society A: Mathematical, Physical & Engineering Sciences*, 457(2016), 2921–2934, 2001.

[16] O.P. Bruno and F. Reitich, Solution of a boundary value problem for the Helmholtz equation via variation of the boundary into the complex domain, *Proceedings of the Royal Society of Edinburgh, Section A, Mathematics*, 122A, 317–340, 1992.

[17] O.P. Bruno, A. Sei and M. Caponi, High-order high-frequency solutions of rough surface scattering problems, *Radio Science*, 37(4), 2-1–2-13, 2002.

[18] M. Cadilhac, Some mathematical aspects of the grating theory, In *Electromagnetic Theory of Gratings*, R. Petit, Ed., Berlin, Springer-Verlag, 1980.

[19] D. Colton and R. Kress, *Integral Equation Methods in Scattering Theory*, New York, John Wiley and Sons, 1983.

[20] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Berlin, Springer-Verlag, 1998.

[21] J. DeSanto, G. Erdmann, W. Hereman and M. Misra, Theoretical and computational aspects of scattering from rough surfaces: one-dimensional perfectly reflecting surfaces, *Waves in Random Media*, 8(4), 385–414, 1998.

[22] A. Dutt and V. Rokhlin. Fast Fourier transforms for nonequispaced data, *SIAM Journal on Scientific Computing*, 14(6), 1368–1393, 1993.

[23] T. Elfouhaily, B. Chapron, K. Katsaros and D. Vandemark, A unified directional spectrum for long and short wind-driven waves, *Journal of Geophysical Research*, 102(C7), 15781–15796, 1997.

[24] Federal Aviation Administration, Navigation Services, GPS Policy-Modernization, *http://www.faa.gov/about/office_org/headquarters_offices/ato/service_units/techops/navservices/gnss/gps/policy/modernization/index.cfm* .

[25] G.B. Folland, *Real Analysis: Modern Techniques and Their Applications*, New York, John Wiley and Sons, 1999.

[26] W. Franz, Über die Greenschen Funktionen des Zylinders und der Kugel, *Zeitschrift für Naturforschung*, 9a, 705–716, 1954.

[27] I. Fredholm, Sur une classe d'équations fonctionnelles, *Acta Mathematica*, 27(1), 365–390, 1903.

[28] P.R. Garabedian, *Partial Differential Equations*, Providence, AMS Chelsea Publishing, 1998.

[29] E. Giladi and J.B. Keller, A hybrid numerical asymptotic method for scattering problems, *Journal of Computational Physics*, 174(1), 226–247, 2001.

[30] M.C. Haslam and O.P. Bruno, High-order solution of the grating diffraction problem for singular one-dimensional perfect conductors, (*in process*).

[31] R.M. James, A contribution to scattering calculation for small wavelengths—the high frequency panel method, *IEEE Transactions on Antennas and Propagation*, 38(10), 1625–1630, 1990.

[32] J.B. Keller, A Geometrical Theory of Diffraction, in *Calculus of Variations and its Applications, Proceedings of Symposia in Applied Math*, 8, 27–52, New York, McGraw-Hill, 1958.

[33] J.B. Keller, Geometric Theory of Diffraction, *Journal of the Optical Society of America*, 52(2), 116–130, 1962.

[34] J.B. Keller and R.M Lewis, Asymptotic methods for partial differential equations: the reduced wave equation and Maxwell's equations, *Surveys in Applied Mathematics, Vol. 1*, J.B. Keller, D.W. McLaughlin and G.C. Papanicolaou, Ed., 1–82, New York, Plenum Press, 1995.

[35] R. Kress, *Linear Integral Equations*, New York, Springer-Verlag, 1999.

[36] H.C. Ku, R.S. Awadallah, R.L. McDonald and N.E. Woods, Fast and accurate algorithm for electromagnetic scattering from 1-D dielectric ocean surfaces, *IEEE Transactions on Antennas and Propagation*, 54(8), 2381–2391, 2006.

[37] R. Kussmaul, Ein numerisches Verfahren zur Lösung des Neumannschen Aussenraumproblems für die Helmholtzsche Schwingungsgleichung, *Computing* 4(3), 246–273, 1969.

[38] R.T. Ling, Numerical solution for the scattering of sound waves by a circular cylinder, *AIAA Journal*, 25(4), 560–566, 1987.

[39] C.M. Linton, The Green's function for the two-dimensional Helmholtz equation in periodic domains, *Journal of Engineering Mathematics*, 33, 377–402, 1998.

[40] E. Martensen, Über eine Methode zum räumlichen Neumannschen Problem mit einer Anwendung für torusartige Berandungen, *Acta Mathematica*, 109(1), 75–135, 1963.

[41] J.R. Mautz and R.F. Harrington, A combined-source solution for radiation and scattering from a perfectly conducting body, *IEEE Transactions on Antennas and Propagation*, 27(4), 445–454, 1979.

[42] D. Maystre, Rigorous vector theories of diffraction gratings, In *Progress in Optics: Volume XXI*, E. Wolf, Ed., Amsterdam, North-Holland Physics Publishing, 1984.

[43] W.C.H. McLean, *Strongly Elliptic Systems and Boundary Integral Equations*, Cambridge, Cambridge University Press, 2000.

[44] A. Meier, T. Arens, S.N. Chandler-Wilde and A. Kirsch, A Nyström method for a class of integral equations on the real line with applications to scattering by diffraction gratings and rough surfaces, *Journal of Integral Equations and Applications*, 12(3), 281–321, 2000.

[45] R.B. Melrose and M.E. Taylor, Near peak scattering and the corrected Kirchhoff approximation for a convex obstacle, *Advances in Mathematics*, 55(3), 242–315, 1985.

[46] L. Nirenberg, Remarks on strongly elliptic partial differential equations, *Communications on Pure and Applied Mathematics*, 8(4), 648–674, 1955.

[47] E. Perrey-Debain, J. Trevelyan and P. Bettess, Use of wave boundary elements for acoustic computations, *Journal of Computational Acoustics*, 11(2), 305–321, 2003.

[48] R. Petit, A tutorial introduction, In *Electromagnetic Theory of Gratings*, R. Petit, Ed., Berlin, Springer-Verlag, 1980.

[49] Lord Rayleigh, On the dynamical theory of gratings, *Proceedings of the Royal Society of London, Series A, Containing Papers of a Mathematical and Physical Character*, 79(532), 399–416, 1907.

[50] Lord Rayleigh, Note on the remarkable case of diffraction spectra described by Prof. Wood, *Philosophical Magazine*, 14, 60–65, 1907.

[51] F. Riesz, Über lineare Funktionalgleichungen, *Acta Mathematica*, 41(1), 71–98, 1916.

[52] V. Rokhlin, Rapid solution of integral equations of scattering theory in two dimensions, *Journal of Computuational Physics*, 86(2), 414–439, 1990.

[53] Y. Saad and M.H. Schultz, GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM Journal of Scientific and Statistical Computing*, 7(3), 856–869, 1986.

[54] M.E. Veysoglu, H.A. Yueh, R.T. Shin and J.A. Kong, Polarimetric passive remote sensing of periodic surfaces, *Journal of Electromagnetics Waves and Applications*, 5(3), 267–280, 1991.

[55] A.G. Voronovich, *Wave Scattering From Rough Surfaces*, New York, Springer-Verlag, 1994.

[56] K.F. Warnick and W.C. Chew, Numerical simulation methods for rough surface scattering, *Waves in Random Media*, 11, R1–R30, 2001.

[57] G.N. Watson, *A Treatise On The Theory Of Bessel Functions*, Cambridge, Cambridge University Press, 1952.

[58] G.B. Whitham, *Linear and Nonlinear Waves*, New York, John Wiley and Sons, 1999.

[59] R.W. Wood, On a remarkable case of uneven distribution of light in a diffraction grating spectrum, *Philosophical Magazine*, 4, 396–402, 1902.

[60] Z. Zhao, L. Li, J. Smith and L. Carin, Analysis of scattering from very large three-dimensional rough surfaces using MLFMM and ray-based analyses, *IEEE Antennas and Propagation Magazine*, 47(3), 20–30, 2005.