Numerical and asymptotic methods for scattering by penetrable obstacles

Samuel Groth

Department of Mathematics and Statistics

University of Reading

A thesis submitted for the degree of

Doctor of Philosophy

February, 2016

Abstract

In this thesis we develop a hybrid numerical-asymptotic boundary element method for the efficient approximation of high-frequency time-harmonic wave scattering by two-dimensional penetrable convex polygons.

It is well known that the numerical simulation of wave scattering problems is computationally intensive when the wavelength is small relative to the scattering obstacle. In fact, conventional methods based on polynomial approximation spaces suffer from the requirement that (in two dimensions) the number of degrees of freedom must grow at least linearly with respect to the frequency in order to maintain a fixed accuracy. The hybrid numerical-asymptotic (HNA) approach aims to remove this restriction by enriching the approximation space with oscillatory basis functions chosen to efficiently capture the high-frequency asymptotics of the solution. Boundary element methods based on HNA approximation spaces have been developed for scattering by different kinds of *impenetrable* obstacle. They have been shown to achieve approximations of a prescribed accuracy with a number of degrees of freedom which grows only very mildly with the frequency and in many cases can remain fixed. This thesis extends the HNA approach for the first time to the *penetrable* case.

More specifically, the HNA method developed here involves first computing the geometrical optics approximation on the boundary of the scatterer using a beam tracing algorithm and then approximating the remaining diffracted field using a boundary element method with an approximation space enriched with oscillatory basis functions. These basis functions are chosen via a heuristic extension of the Geometrical Theory of Diffraction (developed for impenetrable problems) to the penetrable case. We show via numerical experiments that a significant improvement over geometrical optics can be achieved using an approximation space containing only a small number of degrees of freedom. Moreover, as frequency increases we can maintain (or even improve) accuracy without needing to increase the number of degrees of freedom.

Declaration

I confirm that this is my own work and that the use of all material from other sources has been properly and fully acknowledged.

Acknowledgements

Contents

Chapter 1

Introduction

In this thesis, we are concerned with the scattering and absorption of time-harmonic electromagnetic and acoustic waves by penetrable scatterers, that is, regions in which the wave speed differs from that of the background medium. Such a scenario arises in numerous applications of mathematical and physical interest. For example, problems associated with the scattering of laser beams by dielectric metals, ultrasound by organs within the human body, and seismic waves by geological features in the Earth's subsurface, each command active areas of research. We shall focus on one particular application, namely the scattering of light by the ice crystals composing cirrus clouds in the upper troposphere. This application is of significance in climate modelling, specifically in estimating the influence of cirrus on the Earth-atmosphere radiation balance [8].

Conventional numerical methods for solving such wave scattering problems include the finite element method (FEM) and the boundary element method (BEM). When implemented using piecewise polynomial approximation spaces, these methods suffer from the drawback that a fixed number of degrees of freedom is required per wavelength in order to represent the oscillatory solution. This can lead to prohibitive computational expense when the scatterer is large relative to the wavelength, as is often the case in applications. In this "high-frequency" regime one can alternatively appeal to asymptotic approximation techniques such as Geometrical Optics (GO), the Kirchhoff Approximation (KA, sometimes called "Physical Optics") and the Geometrical Theory of Diffraction (GTD). However, although such approximations have a low (and often frequency-independent) computational cost, the price one pays is that they are only accurate for "sufficiently high" frequencies. The question of how high the frequency needs to be for "sufficient accuracy" depends on the particular scattering problem being considered, and moreover is usually not known a priori. In many applications (in particular for the example of light scattering by atmospheric ice crystals mentioned above) there is a significant and important range of frequencies for

which neither conventional numerical methods nor asymptotic methods give satisfactory results.

The *hybrid numerical-asymptotic* (HNA) approach is a general methodology for scattering problems which aims to fuse conventional numerical methods with high-frequency asymptotics to create algorithms that are controllably accurate and computationally feasible over the whole frequency range. The key idea is to enrich the boundary element approximation space with oscillatory functions, chosen using partial knowledge of the high frequency asymptotic behaviour of the solution. More explicitly, one seeks to approximate the unknown solution v of the relevant boundary integral equation using an ansatz of the form

$$v(\mathbf{x}, k) \approx v_0(\mathbf{x}, k) + \sum_{m=1}^{M} v_m(\mathbf{x}, k) \exp(\mathrm{i} k \psi_m(\mathbf{x})), \quad \mathbf{x} \in \Gamma, \qquad (1.1)$$

where k (the wavenumber) is proportional to the frequency of the waves, and Γ is the boundary of the scatterer. In this representation, v_0 is a known (generally oscillatory) function, typically the GO approximation, the phases ψ_m are chosen a priori (derived from the high-frequency asymptotics of the diffracted field) and the amplitudes v_m are approximated numerically using piecewise polynomials. The expectation is that if

(i) v_0, the GO approximation, is calculated accurately, and

(ii) ψ_m are chosen wisely via an understanding of the high-frequency asymptotics of the diffracted field,

then $v_m(\cdot, k)$ will be much less oscillatory than $v(\cdot, k)$ and so can be more efficiently approximated by piecewise polynomials than v itself.

For a number of important classes of scattering problems the HNA approach has been shown to provide a dramatic reduction in the number of degrees of freedom required at high frequencies compared to conventional methods [28]. However, to date, the HNA approach has been applied exclusively to problems of scattering by *impenetrable* scatterers, i.e., where perfectly-conducting, sound-soft (Dirichlet), sound-hard (Neumann) or impedance (Robin) boundary conditions are imposed on the boundary Γ. Moreover, until very recently in [29] (where a class of simple non-convex impenetrable scatterers was considered), its successful application was restricted to convex impenetrable scatterers, for which multiple re-reflections and questions of partial illumination need not be considered.

The main aim of this thesis is to begin the challenging task of generalising the HNA methodology to the case of so-called "transmission problems" for *penetrable* scatterers, where the scatterer is a region in which the wave speed differs from that of the background propagation medium. The main difficulty in the penetrable case is that the high-frequency

2

asymptotic behaviour is significantly more complicated than in the impenetrable case. In particular, the boundary of the scatterer represents the interface between two media with different wave speeds, and hence two different wavenumbers, and so we expect to need to modify the ansatz (1.1) to include terms oscillating with both of these wavenumbers. In addition to the phenomena of reflection and diffraction that occur in the impenetrable case, in the penetrable case we observe a new phenomenon, *refraction*, which occurs when a wave propagating in the exterior medium is transmitted into the scatterer and vice versa. One key difficulty this presents is that a wave propagating inside the scatterer can undergo multiple (in fact, infinitely many) internal reflections/diffractions (this is described in more detail in §5.2). We therefore expect that, in order for the amplitudes v_m to be non-oscillatory, we would need to consider infinitely many different phases ψ_m. (This is in contrast to the case of scattering by sound-soft convex polygons considered in [30] and [63], where the high-frequency behaviour can be completely captured using just two phase functions, i.e., $M = 2$ in (1.1), corresponding to waves travelling clockwise and anticlockwise around the boundary.) This complicates the development of an ansatz of the form (1.1) for the transmission problem, because to create a viable numerical algorithm we must choose only a finite number of these phases.

Another key difficulty is that the high-frequency asymptotic theory for penetrable scatterers is not nearly as well understood as for the impenetrable case. In particular, there is no known closed-form analytical (or even asymptotic) solution to the canonical problem of diffraction by a penetrable wedge, despite many attempts to derive one (see, e.g., [2], [24], [104], [121]). This means that we do not have a fully-developed GTD for penetrable scatterers from which to infer the correct choice of phases ψ_m in the HNA ansatz (1.1).

That being said, one might imagine that the phases describing the diffracted fields in the penetrable case can be determined by heuristically extending the principles of the GTD for the impenetrable case. Some evidence exists in the literature (e.g., [49, 77]) to support the validity of this extension. And we may also look to the time-domain version of the penetrable wedge problem (§4.8.1) to glean further information about the phase structure of the diffracted wave in the interior of the wedge, such as the presence of the so-called head or lateral wave.

Although we may be able to determine with some confidence the phase structure of the diffracted field for the infinite wedge problem, the problem we really wish to solve is for a truncated and closed-off wedge, i.e., a polygon. As mentioned earlier, the diffracted waves will be internally reflected within the polygon giving rise to infinitely many different phase functions ψ_m. One must then confront the question of how many of these phase functions

3

should be included in the HNA approximation space. In the case when the scatterer possesses some absorption (corresponding to the internal wavenumber containing a positive imaginary component), which is often the case in applications (see §1.1), the waves propagating within the polygon decay exponentially with distance. Hence, we shall expect that we may only require to include a small number of these phase functions. In fact, in the high-absorption limit, the penetrable scattering problem can then be well-approximated by the impenetrable scattering problem with impedance boundary conditions (see, e.g., [57] or [3]) for which the HNA ansatz (1.1) with $M = 2$ is appropriate [31]. As the level of absorption is reduced, the waves propagating within the polygon become more prominent and we must include extra phases in the ansatz (1.1).

We aim to construct simple modifications to the ansatz (1.1) which efficiently approximate solutions to transmission problems. To this end, we shall endeavour to include phase functions corresponding to the leading order components of the diffracted field and ignore those corresponding to higher-order diffracted terms. We shall see that this approach is particularly effective when the scatterer possesses some absorption. Nevertheless, even with zero absorption, considering only a small number of the leading order components can lead to an effective HNA BEM since here the diffracted waves decay as they travel, albeit at a slower rate. The process of constructing appropriate versions of the HNA ansatz shall be detailed in §5.

In anticipation of later discussions, we pause here to make a brief technical remark. The evaluation of the GO term v_0 is considerably more difficult when absorption is present. This difficulty stems from the fact that, since the interior wavenumber is complex, the propagation direction vectors associated with the geometrical optics waves (or beams) within the polygon are also complex. Making sense of such direction vectors has been studied before (see, e.g., [32, 33, 116]), however none of these studies have both considered the precise problem of interest here and compared the obtained approximation to an "exact" solution. In §4.3 we employ a method for calculating the GO approximation similar to some previous authors (in particular, [116] and [32]), however we observe a phenomenon which appears to be unreported in the literature and dramatically affects the approximation at angles associated with total internal reflection. In fact, at these angles, the GO approximation is almost completely out of phase with the exact solution. An investigation is carried out in §4.3.4 which identifies the occurrence of this phenomenon and shows that a different sign choice in one of the equations acts to remedy the situation. There remains, nevertheless, a small (and shrinking as $k \to \infty$) transition region where neither sign choice gives a good approximation and further work, not performed here, is required to improve the approximation in this region.
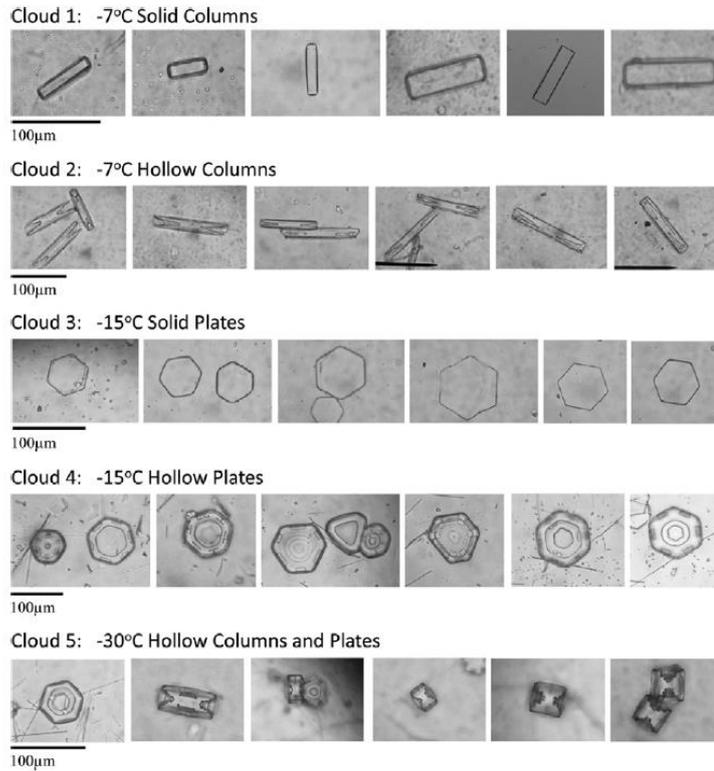
Figure 1.1: Some typical ice crystal geometries: solid and hollow hexagonal plates and columns. Image taken from [125].

1.1 Background and motivation

Before commencing with a thorough outline of the thesis, we first present some background into the problem of light scattering by ice crystals to provide motivation for the development of the HNA BEM of this thesis.

It is well established that understanding the light scattering properties of atmospheric ice crystals is important in modelling the radiation balance of cirrus clouds [7,8,92]. Due to the wide coverage of cirrus over the Earth (approximately 30% at any one time in the mid-latitudes, and 60-80% in the tropics [7,92,123]), these clouds in turn play an important role in the Earth-atmosphere radiation balance. Therefore methods which can provide accurate approximations for scattering by these ice crystals are extremely valuable to atmosphere and climate scientists. Many such methods exist however no single method is, as yet, applicable to all of the myriad scattering problems arising within cirrus clouds. We now consider the variations that lead to the zoo of combinations of particle geometry, incident radiation wavelength and particle refractive index.

The ice crystals within cirrus exhibit a large array of sizes and shapes [48,64,89]. These

5

Figure 1.2: A bullet-rosette formed from the aggregation of hexagonal columns. Image taken from [91].

shapes are mostly variations on the hexagonal column or plate as shown in Figure 1.1. In this figure we observe the most commonly occuring ice crystal shape, namely the "solid" hexagonal column which is called a "plate" when the length is the column is short compared to its hexagonal face. This figure also shows variations on the solid column which commonly occur; these are the so-called "hollow" columns and plates which are non-convex owing to inclusions at either end. These hexagonal columns also often aggregate together to form "bullet-rosettes" such as that shown in Figure 1.2.

The fact that ice tends to form with a hexagonal structure was first noted by Johannes Kepler [80] and he hypothesised that the hexagonal symmetry was related to the optimal stacking of spheres, which was later found to be a correct analogy [65]. Nevertheless, there often occur in nature so-called trigonal ice crystals which possess a three-fold symmetry [110]. We shall use this as motivation to consider both triangular and hexagonal ice crystal geometries in this thesis.

Ice crystals also vary in size, with a characteristic length a (e.g., the radius of the smallest sphere which circumscribes the crystal) ranging from less than 10 microns to many thousands of microns [138]. The wavelength λ of the radiation incident upon the clouds also ranges from 10^{-7}m (ultraviolet) to 10^{-1}m (microwave). We define the dimensionless *size parameter*

$$\chi := \frac{2\pi a}{\lambda}$$

to describe the size of the ice crystals relative to the wavelength of the incident light. Then we can see that the size parameter range of interest is approximately $10^{-6} < \chi < 10^4$.

Finally, we note that the refractive index of ice varies with the wavelength of the incident radiation. Figure 1.3 shows the real part of the refractive index of ice for electromagnetic radiation with wavelength ranging from $10^{-1}\mu m$ to $10^4\mu m$. We see that it fluctuates

6
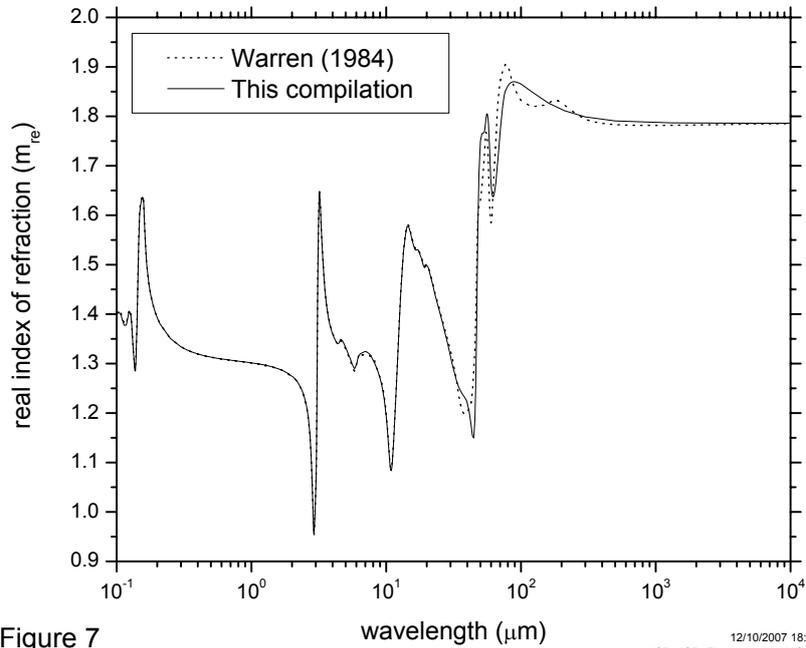
Figure 7

Figure 1.3: Real part of the complex refractive index of ice; comparison of two different compilations of measurements, the first made in 1984 in [139] and the second in [140]. Plot taken from [140].

between a minimum value of approximately 0.95[1] and a maximum value of approximately 1.9. Similarly, the imaginary part of the refractive index varies with wavelength from a minimum of 0 to a maximum of approximately 1.

These variations in particle geometry, relative size, and refractive index combine to create an enormous variety of scattering problems to be solved. Over the years, many methods have been developed for tackling different problems within the myriad combinations of particle shape, size and incident radiation frequency. These methods fall into two main camps:

- numerical (or "exact") methods,

- asymptotic (or "approximate") methods.

---

[1] A refractive index of less than 1 leads to a "phase velocity" greater than the speed of light, although the "signal velocity" is not. See a classical physics textbook, such as [105], for more detail.

7

1.2 Numerical and asymptotic methods for scattering problems

Numerical methods either discretise the underlying equations (Maxwell or Helmholtz) and solve the resultant discrete system, or adopt a truncated separation of variables ansatz and obtain the associated coefficients by numerically enforcing the boundary conditions at the scatterer's surface. These methods are exact in the sense that, with them, one may in principle achieve arbitrary accuracy if the numerical approximation is sufficiently refined. However, they suffer from the fact that to achieve a satisfactory accuracy, each wavelength of the scattered field must be resolved. This leads to an associated computational cost that scales with the frequency of the incident wave. Therefore they are in practice limited to low-frequency problems.

Asymptotic methods, on the other hand, appeal to the high-frequency behaviour of light to justify the implementation of geometrical techniques. These methods have a computational cost which is very low and often independent of the frequency, however they are only accurate for sufficiently high frequencies.

The hybrid numerical-asymptotic approach seeks to unify these two classes of technique by developing algorithms that are numerically "exact" but have a cost which is independent of frequency.

There is in fact a third class of methods which should be briefly mentioned, namely analytical methods. Maxwell's equations and the Helmholtz equation can be solved exactly by the method of separation of variables when the scattering body is a shape for which there exists a coordinate system in which the Helmholtz equation is separable, namely spherical and spheroidal scatterers. The separation of variables solution process is often referrred to as Mie theory, Mie-Lorenz theory or Lorenz-Mie theory in the electromagnetics literature, e.g., [105]. This analytical approach generally results in a solution of the form of an infinite sum, and the number of terms which must be taken for an accurate solution increases drastically with frequency. Hence efficient numerical techniques (such as hybrid methods) can often prove faster at high frequencies than analytical techniques even when an exact solution exists.

In addition to separation of variables, there exist the following closely related analytical techniques: the Kontorovich-Lebedev transform (see, e.g., [113]); the Wiener-Hopf technique (see, e.g., [76]); and the Sommerfeld-Malyuzhinets method [6]. However, these have so far been able to yield exact solutions only for problems involving scattering by infinite lines or wedges with different boundary conditions, although not the transmission boundary conditions as we shall discuss in §4.8. One final technique which can be viewed as

8

analytical is the Unified Transform Method of Fokas and co-authors [128], however this is also limited currently in its applicability for scattering problems, in particular to 2D interior problems.

We will briefly summarise some of the most popular numerical and asymptotic methods for solving wave scattering problems with the aim of placing HNA methods within the context of the most competitive current computational techniques. Figure 1.4 shows the methods we shall discuss and how they relate to one another. For more detailed reviews, the reader is referred to [15, 78, 85, 144].

1.2.1 Numerical methods

Numerical methods can be subdivided roughly into two distinct groups: differential equation techniques, which tackle the governing partial differential equations directly, and integral equation techniques, which first reformulate the boundary value problem as a system of boundary integral equations and then discretise these equations.

Differential equation techniques are the most general in that they are suitable for heterogeneous media (i.e., media with a spatially varying wavenumber) and they are arguably the simplest to implement. The two most popular such techniques are the Finite Difference Time-Domain (FDTD) method [131, 132] and the FEM [75].

The FDTD method, which solves the time-dependent Maxwell's equations, was pioneered by Yee in 1966 [150] and has since become an extremely popular method within the atmospheric sciences community for calculating solutions to scattering problems [146]. Its popularity stems from the simplicity of its implementation. The solution domain is discretised into a rectangular grid and the differential operators are approximated via finite differences. The problem of scattering by a pulse is often considered and the solution is then mapped to the frequency domain using the discrete Fourier transform. The method is very flexible in terms of particle geometry and heterogeneity. However, its main drawbacks are that one must discretise an artificially truncated domain (in the case of an exterior scattering problem) larger than the particle and it suffers from dispersion errors. Also, its demand for memory is high, with the number of degrees of freedom (#DOF) scaling with the size parameter χ worse than $\mathcal{O}(\chi^{d+1})$ where $d = 2, 3$ is the spatial dimension of the problem. (Note that the size of the matrix to be stored scales as $(\#\text{DOF})^2$.)

The FEM solves the frequency-domain boundary value problem by using a variational form. Similar to the FDTD method, the FEM discretises a truncated solution domain but typically with triangular elements rather then rectangular ones. This allows for a much more accurate representation of the particle shape. The FEM can handle complicated particle geometries and can deal with heterogeneity, however it possesses similar drawbacks

to the FDTD method, with its #DOF scaling as $\mathcal{O}(\chi^d)$. However, one attractive feature is that the system matrix is sparsely populated.

Integral equation techniques are popular for scattering problems in which the exterior unbounded medium is homogeneous since they lead to a finite computational domain. This domain is either the surface of the scatterer in the case of boundary integral equations (BIEs) or the volume of the scatterer in the case of volume integral equations (VIEs). We note that BIEs also have a computational domain of reduced dimension, however this requires the scatterer to be homogeneous too. The BEM, also known as the Method of Moments, is the most popular numerical method associated with BIEs. Owing to its reduced dimension its #DOF scales as $\mathcal{O}(\chi^{d-1})$, however it should be noted that the resulting system matrix is densely populated in contrast to that for the FEM. VIEs differ from BIEs in that they are suitable for heterogeneous particles. A popular numerical method based on VIEs is the Discrete Dipole Approximation (DDA) which discretises the volume of the scatterer into a periodic cubic lattice so that the fast-Fourier transform may be employed for calculating the matrix-vector products [151]. The #DOF scales as $\mathcal{O}(\chi^d)$. It should be pointed out that it is not necessary to store the whole matrix of size $(\#DOF)^2$ for BEM when it is used in conjunction with the fast-multipole method [44] or $\mathcal{H}$-matrix compression [56]. This is also true for VIE methods when they are used with the fast-Fourier transform [112]. Such acceleration techniques reduce the storage requirements to $\mathcal{O}(\#DOF \log(\#DOF))$ although the constant associated with this can often be large, and nevertheless the storage requirement still increases as χ increases .

One final method we mention which also stems from BIEs for scattering problems is the Null-Field Method (NFM), also known as the T-matrix method, and is similar to the Method of Fundamental Solutions (MFS) [10]. The NFM dates back to the work of Waterman in the 1960s (see, e.g., [141]) and detailed derivations may be found in [100] for acoustics and [105] for electromagnetics. The NFM begins with the fact that the free-space solution to the underlying equations may be rewritten as an infinite sum over spherical wave functions multiplied by Bessel functions. Upon truncation of this sum and substitution into the BIEs, one obtains a representation for the solution also as a sum over products of spherical wave functions and Bessel functions with coefficients to be determined by integrations of these functions around the boundary. This method is extremely efficient and reproduces the exact solution when the scatterer is either a sphere or spheroidal in shape. There are many drawbacks, however, which include: the fact that the representation for the solution is only valid outside of the smallest sphere enscribing the scatterer; the error associated with truncating the series after N terms is not well understood; the linear system often becomes ill-conditioned as N grows large. The main advantage of this method is that

once the T-matrix has been calculated, scattering by particles in random orientation can be solved for in a single calculation rather than solving for many incident wave directions and averaging the solutions.

1.2.2 Asymptotic methods

The most elementary asymptotic method for wave scattering problems is the geometrical optics (GO) approximation. In GO, the incident wave is decomposed into a bundle of rays or beams which are tracked via the laws of reflection and refraction as they strike the scattering surface/surfaces. Algorithms calculating the GO approximation are called ray tracing or beam tracing algorithms. In practical applications, such algorithms are often employed with the addition of Fraunhofer diffraction in the far-field zone [21,94,109,143], following from Babinet's principle applied to the flat obstacle which corresponds to the particle's geometric shadow area. Babinet's principle states that the diffraction pattern of a flat obstacle is identical to that of an aperture of the same size and shape. Such methods have been termed Conventional Geometrical Optics Methods (CGOM) and have been reported to be accurate for $\chi > 100$ [16]. However, as we shall see in this thesis, such a broad statement is difficult to justify since the accuracy of GO is very much dependent on geometrical parameters such as the number of scattering vertices and the refractive index of the particle. Also, such a statement is very dependent upon the required solution accuracy.

An alternative simple extension to GO is the Kirchhoff approximation (or Physical Optics approximation) which was originally developed in the context of diffraction by apertures [76] (from which Fraunhofer theory also derives). This is obtained via substitution of the GO approximation on the boundary of the scatterer into the boundary integral representation. This method has been well-studied and is known to provide an improvement over CGOM [16, 107] with applicability reported for particles with $\chi > 20$ for the algorithm in [16]. However, in [16] no numerical errors are presented and it appears from far-field comparison plots that the overall accuracy is poor for small χ although some important features in the scattering pattern are discernible. We present results for the Kirchhoff approximation developed for our ice scattering application in §4.6.

A more systematic extension of GO is the Geometrical Theory of Diffraction (GTD), pioneered by Joseph Keller [79], which postulates rays in addition to those of GO. These additional rays emanate from edges and vertices of the scatterer and together compose the diffracted field. GTD specifies the phase and amplitude of these rays by analysing the solutions of relevant canonical problems, such as scattering by an infinite half-plane or an infinite wedge, for which the asymptotic behaviour of the scattered field can be obtained. Therefore, these methods are limited in their scope to those problems for which the relevant

canonical problems can be solved exactly (or at least asymptotically). In particular, they have not been successfully applied to transmission problems owing to a lack of a solution to the canonical problem of scattering by a penetrable wedge (see §4.8). We note also that the GTD has been extended via employment of the boundary integral representation (in much the same way as the Kirchhoff approximation is obtained from GO) to the Physical Theory of Diffraction (see [136]), however this theory suffers from the same aforementioned drawbacks as the GTD.

1.3   The hybrid numerical-asymptotic method

An in-depth historical review of the hybrid numerical asymptotic method may be found in [28] and so here we shall merely outline some significant milestones in the development of the method so as to provide some context for the new developments in this thesis.

The first appearance of the hybrid numerical-asymptotic approach in print was in a paper by Uncles in 1976 [137] in which the numerical solution of "moderate-frequency" scattering by a sound-hard sphere was discussed. In that paper, Uncles reformulated the scattering problem as an integral equation on the boundary of the scattterer and made the ansatz

$$v = V \exp\left(\mathrm{i}k\mathbf{d}^i \cdot \mathbf{x}\right), \quad \mathbf{x} \in \Gamma, \tag{1.2}$$

for the unknown boundary data, where $\mathbf{d}^i$ is the direction of the incident wave. Uncles showed, via numerical examples, that the function V could be approximated much more efficiently (i.e., on a coarser computational mesh) than v itself.

This brief paper ignores the fact that the simple ansatz (1.2) breaks down in regions (so-called "Fock domains") of size $\mathcal{O}(k^{-1/3})$ around shadow boundaries and that it neglects creeping rays. It also does not discuss the efficient evaluation of the oscillatory integrals that arise as a result of employing the ansatz (1.2) on a coarse mesh. These issues were addressed to some extent by Bruno et al. [23] where the mesh was adapted to be finer in the Fock domains to deal with the breakdown of (1.2). The arising oscillatory integrals were evaluated using the method of stationary phase. More rigorous HNA methods for this problem have been put forth in [40] and [71] in which the full asymptotic behaviour of the solution of the boundary is taken into account. In [40] it is shown rigorously that the proposed method achieves an error that depends only very mildly on the frequency k. The oscillatory integrals in [40] were later dealt with efficiently via a Filon method in [81] and those in [71] were handled via the method of Numerical Steepest Descent (NSD).

Of more relevance to the scattering problem of interest in this thesis are the applications of the HNA approach to polygonal scatterers, for which an ansatz of the form (1.1) is

required. The first polygon problem to be tackled effectively was scattering by a sound-soft convex polygon in 2007 by Chandler-Wilde and Langdon [30]. There the authors proposed the ansatz

$$v = v_{go} + v^+(s)\exp(\mathrm{i}ks) + v^-(s)\exp(-\mathrm{i}ks), \qquad (1.3)$$

for the unknown boundary data on a typical side of the polygon. Here v_{go} is the geometrical optics approximation to the solution and s is the arc-length along the side. The second and third terms on the right hand side correspond to diffracted waves travelling in both directions around the boundary of the scatterer. In [30] an HNA Galerkin h-BEM based on this ansatz was proposed and analysed. It was proved that in the ansatz (1.3) the functions $v^+(s)$ and $v^-(s)$ are non-oscillatory and that they may be approximated using a number of degrees of freedom which is required to grow only logarithmically as a function of k in order to achieve a prescribed level of accuracy. The same ideas were used to propose an HNA Galerkin hp-BEM for the same problem in 2013 [63] and were subsequently extended to convex polygons with impedance boundary conditions [31] and non-convex polygons [29]. For each of these problems, the same logarithmic dependency of the error on the wavenumber k was shown.

Remark 1.1 (A note on nomenclature). *The term "hybrid numerical-asymptotic" appears to have first been coined in this context in the paper [51] by Giladi and Keller in 2001. However, the method proposed there was a FEM rather than a BEM. Although the same authors extended these ideas to boundary integral equations in [50], they did not reuse the HNA name. The first application of the name within a BEM context was in [40] in 2007. Nevertheless, numerous different titles have been applied to the HNA BEM methods described above. These include "$\mathcal{O}(1)$ method" [23], "high-frequency BEM" [63] and "phase extraction technique" [71]. Crucially, all these different methods employ an ansatz of the form (1.1) within a BIE setting and so fall beneath (in the terminology employed in this thesis) the HNA umbrella.*

Also, it must be noted that there are some related but crucially different methods going by similar names. For example the term "hybrid asymptotic-numerical" has been used by Barbone and Michael in [9] to describe a method which employs a BEM in the vicinity of diffraction points (e.g., vertices) and matches the numerically approximated solution there to an asymptotically derived solution away from these points.
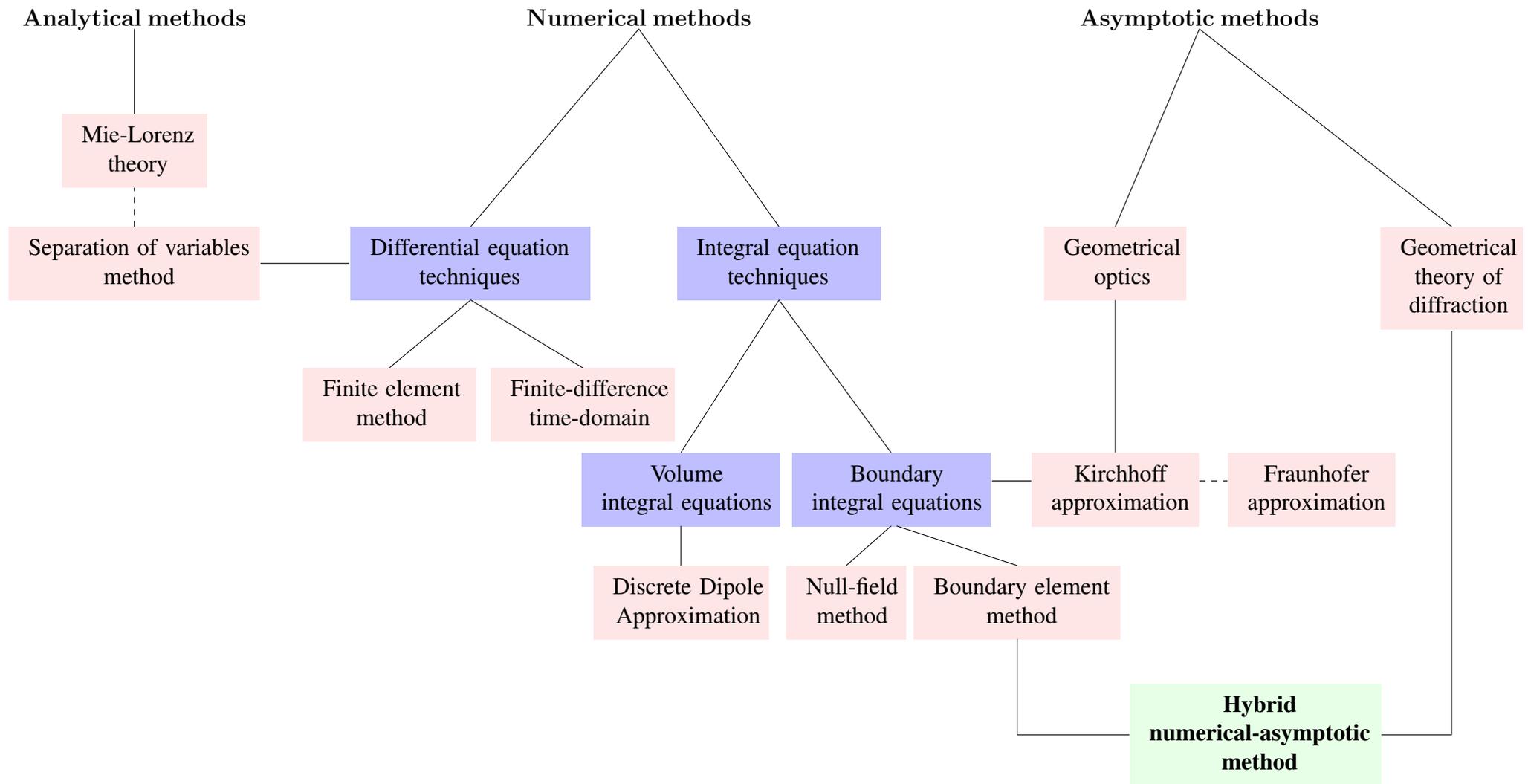
Figure 1.4: Relationships between popular methods for wave scattering problems.

1.4 Main aims and key results

The main aim of this thesis is to extend the HNA approach to scattering by 2D penetrable convex polygons in order to develop a numerical method that can produce approximations of an accuracy sufficient for many applications without the requirement that the number of degrees of freedom in its approximation space must increase with frequency in order to maintain this accuracy. We shall refer to this sufficient level of accuracy as "engineering accuracy" and we note that this is a common goal in calculations for wave problems [47, 87, 98]. A specific numerical value for this accuracy seems to never be given but it may be interpreted to mean of the order of a few percent, or such that the solution "looks correct". In this thesis, many of the calculations shall aim to achieve approximately 1% relative L^2 error in the solution away from the scatterer boundary or in the far-field. We can be confident that such an accuracy is equivalent to, if not better than, "engineering accuracy". It has been reported that six to ten degrees of freedom per wavelength are usually required with standard methods for this to be maintained [98].

We remark that our aim is slightly different to that in [29, 40, 63] where HNA BEMs for impenetrable scatterers are presented. There the BEMs also have the desirable property that they are controllably accurate for a fixed frequency which is possible since the ansatzes capture all the oscillatory behaviour of the diffracted field. For the penetrable problem, however, the oscillatory behaviour of the diffracted field is considerably more complicated and to include all of it would be impossible. Hence we shall aim to incorporate only the low-order diffracted terms in our HNA ansatz in order to capture the majority of the field with a small number of degrees of freedom. We shall see that, with a BEM based on a simple HNA ansatz constructed in this way for the transmission problem, we can significantly improve on classical asymptotic techniques such as GO and the KA using just a small number of degrees of freedom which does not increase with increasing frequency. Indeed, we shall see that, when a small amount of absorption is present in the scatterer, we may achieve engineering accuracy.

The aim to develop such an HNA BEM for the transmission problem may be broken down into the following smaller goals:

- develop an accurate method to calculate the geometrical optics approximation to scattering by absorbing (and non-absorbing) 2D convex polygons,

- develop an HNA ansatz for the penetrable scattering problem,

- implement an approximation space based on this ansatz within a Galerkin hp-BEM.

The novel contributions of this thesis in the pursuit of these goals are the following.

1. A previously unreported phenomenon associated with plane waves propagating between two absorbing media is observed in §4.3.4. We show that ignoring this phenomenon leads to a seemingly erroneous phase in the solution at particular incident angles. We present a remedy to correct for this.

2. A beam-tracing algorithm (BTA) based on this novel interpretation of the planar interface problem is presented. We implement a Kirchhoff approximation which employs this BTA in §4.6.

3. The development and verification of an HNA ansatz for the transmission problem (§5). The verification is performed by computing numerical best approximations (from the HNA approximation space) to reference solutions calculated using a high-order conventional BEM. We show that a prescribed error tolerance may be achieved with a number of degrees of freedom which is small and not increasing with increasing wavenumber k. Some of these results have been published in [54].

4. The implementation of this HNA ansatz within a Galerkin hp-BEM (§6). We present results demonstrating that this BEM produces approximations close to the numerical best approximations.

A complementary goal of this PhD research was to investigate the effectiveness of BEM solves for 3D electromagnetic (EM) scattering problems of relevance in climate science. The main contribution in this regard is the following.

5. The first detailed application of the boundary element method to the 3D EM transmission problem of light scattering by ice crystals of complex shape (§3.3). The boundary element method employed is the open-source software BEM++ [124]. Here we shall only present the results from this study that are pertinent to the topic of this thesis. A more detailed set of results can be found in the publication [53].

1.5 Outline of the thesis

In §2 we describe the boundary value problems we study in this thesis and their reformulations as systems of boundary integral equations. First, we consider the 3D EM problem of scattering by a dielectric particle. Then we review the fact that in 2D, the EM problem reduces to solving two 2D acoustic (i.e., scalar) problems. Since the 2D acoustic case is the main focus of the thesis, we describe its associated BIE reformulation in more detail than

that for the 3D EM problem and provide a self-contained proof that the resulting integral equations possess a unique solution. Finally, we describe some scattering properties which are of interest in applications.

In §3 we discuss the conventional boundary element method and detail the implementation of it in Galerkin form in the 2D case. This includes an exposition of efficient quadrature techniques for the evaluation of the arising 2D weakly singular integrals. The 2D Galerkin hp-BEM described here will serve to provide reference solutions for use later in the thesis. In §3.3 we briefly present some details and results of the application of the open-source boundary element library BEM++ to the 3D EM problem, described more fully in [53].

§4 examines the high-frequency asymptotic behaviour of the boundary solution in the 2D case. In particular, we develop the beam-tracing algorithm which shall be used to calculate the GO term v_0 in (1.1). The BTA relies on the canonical problem of a plane wave at an interface between two absorbing media which we study in detail in §4.3. Later in the chapter, we analyse the asymptotic behaviour of the diffracted field in order to derive a suitable version of the ansatz (1.1) for our transmission problem. This involves a brief exposition of the Geometrical Theory of Diffraction and a study of the canonical problems of the scattering of a plane wave by a penetrable wedge and a point source above a planar interface.

In §5 we develop an HNA boundary element approximation space for our scattering problem. This is a more sophisticated approximation space than that presented in our publication [54]. We demonstrate via numerical best-fitting to reference solutions that this approximation space allows us to obtain accurate and frequency-independent (in terms of number of degrees of freedom) approximations for scatterers with some absorption. Even when no absorption is present, we still see that with a small and frequency-independent number of degrees of freedom, we can achieve a significant improvement in accuracy compared to employing the GO approximation on its own.

§6 details the implementation of a Galerkin hp-BEM which utilises this HNA approximation space. Such an implementation is non-trivial and requires the evaluation of highly-oscillatory integrals as well as careful normalisation of particular entries in the matrix to avoid conditioning issues. It is shown via numerical experiments that the HNA BEM achieves close to the best approximations and is applicable to different convex polygons.

We conclude the thesis in §7 with a brief synopsis of our key achievements and provide ideas for the future development of this work.

Chapter 2

Problem formulation

This chapter is dedicated to the statement of the 3D electromagnetic and 2D acoustic transmission boundary value problems (BVPs) along with their reformulations as systems of boundary integral equations (BIEs). The 3D EM problem is not the focus of this thesis, however it is considered here for good reasons: it is the version of the problem that is of primary interest to applications; it reduces to the 2D acoustic problem upon the assumption of the independence of the field from one spatial direction; and, finally, we illustrate some key features of the conventional BEM via application to the 3D EM problem, in §3.3. We shall be more brief in our presentation of the 3D EM problem than of the 2D acoustic problem since the latter is of primary interest in this thesis. In particular, we shall not prove the unique solvability of the BIE system for the EM problem whereas we shall do so for the 2D problem.

We begin in §2.1 by stating the governing equations of electromagnetism - Maxwell's equations - and in §2.2 their form for time-harmonic fields. This allows us to give in §2.3 the precise statement of the 3D EM scattering problem to be solved. It is then shown how this BVP may be reformulated as a uniquely solvable system of BIEs. The BIE formulation chosen here is the popular *combined-field formulation* which is a well-studied formulation [102].

In §2.5 we review the fact that when the dependence of the electric and magnetic fields is reduced by one dimension, the EM scattering problem is equivalent to two 2D scalar (or acoustic) transmission problems, one for each of the electric and magnetic fields. We state the general transmission problem for the 2D case in §2.6 and reformulate it as a system of BIEs in §2.7. The BIEs we derive are similar to those in [35, 83] (where only smooth scatterers are considered), and also to those proposed in [133] (albeit for an indirect method, in which the unknowns are non-physical "densities", rather than the boundary data itself as is the case here). We remark that other BIE formulations for the transmission problem are also possible - see, e.g., [37, 66, 84, 88, 118]. The unique solvability of our system of BIEs is

known for smooth scatterers [35], however its proof for non-smooth (Lipschitz) scatterers does not appear to be present as a whole in the literature. Therefore, for completeness, we provide such a proof here.

We conclude in §2.8 by detailing briefly the scattering properties which are often of interest in applications. These constitute the scattered far-field and quantities derived from it, such as the amplitude scattering matrix.

2.1 Maxwell's equations

The electric vector $\mathcal{E}$ and the magnetic vector $\mathcal{H}$ together constitute the electromagnetic field. Their behaviour is encapsulated in the four Maxwell's equations which are, assuming the medium is homogeneous and isotropic [18],

$$\nabla \times \mathcal{H} - \varepsilon_0 \frac{\partial \mathcal{E}}{\partial t} = \sigma \mathcal{E} \tag{2.1}$$

$$\nabla \times \mathcal{E} + \mu \frac{\partial \mathcal{H}}{\partial t} = 0 \tag{2.2}$$

$$\nabla \cdot \mathcal{E} = \frac{\rho}{\varepsilon_0} \tag{2.3}$$

$$\nabla \cdot \mathcal{H} = 0, \tag{2.4}$$

where ε_0 is the dielectric constant, σ is the electrical conductivity, μ is the magnetic permeability, and ρ is the electric charge density.

For non-magnetic, transparent substances such as air and ice, μ is equal to unity. For generality, however, we retain μ in all of the formulae in this chapter but we are mostly interested in the case $\mu = 1$. The electrical conductivity σ affects the rate of decay of an electromagnetic field in a medium and contributes to the imaginary part of the medium's refractive index, as we shall see shortly. The electrical charge density is set as

$$\rho = 0$$

since we assume there are no electric charges in the regions concerned.

We have assumed the medium is isotropic and homogeneous so that ε_0, σ and μ are constant. However, we are interested in the case where two differing media are present, each of which is characterised by its own values of these constants. At the interface between the two media, the values of ε_0, σ and μ change discontinuously. Therefore we must impose appropriate "jump" conditions on $\mathcal{E}$ and $\mathcal{H}$ at this interface.

The appropriate boundary conditions can be found by replacing the plane of discontinuity (i.e. the interface between the two media) with a thin layer in which ε_0, σ and μ vary

quickly but continuously between the two media, and then taking limits as the boundary is approached (see [19, Chapter 1] for more details). The results of this procedure are the conditions

$$\mathbf{n} \times [\boldsymbol{\mathcal{H}}] = 0, \tag{2.5}$$

$$\mathbf{n} \times [\boldsymbol{\mathcal{E}}] = 0, \tag{2.6}$$

$$\mathbf{n} \cdot [\mu \boldsymbol{\mathcal{H}}] = 0, \tag{2.7}$$

$$\mathbf{n} \cdot [\varepsilon_0 \boldsymbol{\mathcal{E}}] = 0, \tag{2.8}$$

where $[\cdot]$ denotes the jump of a quantity across the interface and $\mathbf{n}$ is the vector normal to the interface. Physically, (2.5)-(2.6) state that the tangential components of $\boldsymbol{\mathcal{H}}$ and $\boldsymbol{\mathcal{E}}$ must be continuous across the boundary, and (2.7)-(2.8) state that the normal components of $\mu \boldsymbol{\mathcal{H}}$ and $\varepsilon_0 \boldsymbol{\mathcal{E}}$ must be continuous across the boundary. In our case, however, we are assuming $\mu = 1$ throughout and ε_0 to be constant throughout so these two constants may be removed from these jump relations.

2.2 Time-harmonic fields

We now further specialise our formulae by restricting ourselves to time-harmonic fields, i.e., those which have a time-dependence $\mathrm{e}^{-\mathrm{i}\omega t}$. Such EM fields are of the form

$$\boldsymbol{\mathcal{E}}(\mathbf{x}, t) = \mathrm{Re}\{\mathbf{E}(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}\}, \quad \boldsymbol{\mathcal{H}}(\mathbf{x}, t) = \mathrm{Re}\{\mathbf{H}(\mathbf{x})\mathrm{e}^{-\mathrm{i}\omega t}\}. \tag{2.9}$$

Upon substitution of (2.9) into the Maxwell equations (2.1)–(2.4), the time-harmonic Maxwell equations arise:

$$\nabla \times \mathbf{H} + \mathrm{i}\omega\varepsilon\mathbf{E} = 0, \tag{2.10}$$

$$\nabla \times \mathbf{E} - \mathrm{i}\omega\mu\mathbf{H} = 0, \tag{2.11}$$

$$\nabla \cdot \mathbf{E} = 0, \tag{2.12}$$

$$\nabla \cdot \mathbf{H} = 0, \tag{2.13}$$

where $\varepsilon = \varepsilon_0 + \mathrm{i}\sigma/\omega$.

Owing to the relationship of $\mathbf{E}$ to $\mathbf{H}$ as seen in (2.10)–(2.11), it is necessary to solve for just one of the fields. To obtain equations for $\mathbf{E}$ and $\mathbf{H}$ separately, we substitute (2.10) into (2.11), and (2.11) into (2.10) to find that $\mathbf{E}$ and $\mathbf{H}$ both satisfy:

$$\nabla \times \nabla \times \mathbf{E} - k^2\mathbf{E} = 0, \tag{2.14}$$

$$\nabla \times \nabla \times \mathbf{H} - k^2\mathbf{H} = 0, \tag{2.15}$$

where we note that they are still supplemented by equations (2.12)–(2.13). The constant $k := \omega\sqrt{\mu\varepsilon}$ is called the *wavenumber* because $k/2\pi$ gives the number of wavelengths per unit length. The factor $\sqrt{\mu\varepsilon}$ is called the *refractive index* of the medium and is denoted by $n := \sqrt{\mu\varepsilon}$. Note that for $\sigma \neq 0$, $n \in \mathbb{C}\backslash\mathbb{R}$. It is helpful for later when we consider two-dimensional EM problems to note that the vector wave equations (2.14)–(2.15) can be written as vector Helmholtz equations. This can be seen by using the vector identity

$$\mathbf{curl\ curl} \equiv \mathbf{grad\ div} - \Delta \tag{2.16}$$

and noting that $\mathbf{E}$ and $\mathbf{H}$ are both divergence free by (2.3) and (2.4), so that we have

$$\Delta\mathbf{E} + k^2\mathbf{E} = 0, \tag{2.17}$$

$$\Delta\mathbf{H} + k^2\mathbf{H} = 0. \tag{2.18}$$

The time-harmonic versions of the jump-relations (2.5)–(2.8) are simply

$$\mathbf{n} \times [\mathbf{H}] = 0, \tag{2.19}$$

$$\mathbf{n} \times [\mathbf{E}] = 0, \tag{2.20}$$

$$\mathbf{n} \cdot [\mathbf{H}] = 0, \tag{2.21}$$

$$\mathbf{n} \cdot [\mathbf{E}] = 0, \tag{2.22}$$

Note that the final two may be neglected since the normal component of $\mathbf{H}$ being continuous across the boundary is equivalent to the tangential component of $\mathbf{E}$ being continuous across the boundary, and similar for the normal component of $\mathbf{E}$ and the tangential component of $\mathbf{H}$. Hence, from now on we shall only use the first two conditions.

Now we are ready to state the electromagnetic transmission problem which we choose (as is traditional in the literature) to formulate for the electric field.

2.3   The electromagnetic transmission problem

Consider the scattering of a monochromatic plane wave with time-dependence $\mathrm{e}^{-\mathrm{i}\omega t}$ by an open and bounded set $\Omega_2 \subset \mathbb{R}^3$. Suppose that Ω_2 is occupied by a homogeneous and isotropic dielectric material (see Figure 2.1) with a complex refractive index $n_2 := \sqrt{\mu_2\varepsilon_2}$. Further it is assumed that Ω_2 is surrounded by an unbounded homogeneous medium $\Omega_1 := \mathbb{R}^3\backslash\overline{\Omega_2}$ with refractive index $n_1 := \sqrt{\mu_1\varepsilon_1}$ where $\overline{\Omega_2}$ denotes the closure of Ω_2 in $\mathbb{R}^3$. Throughout, Ω_2 will have a Lipschitz boundary $\Gamma = \partial\Omega_2$ (for a definition see, e.g., [103, pp. 89–91]). Throughout we shall refer to Ω_1 and Ω_2 as the exterior and interior domains, respectively.
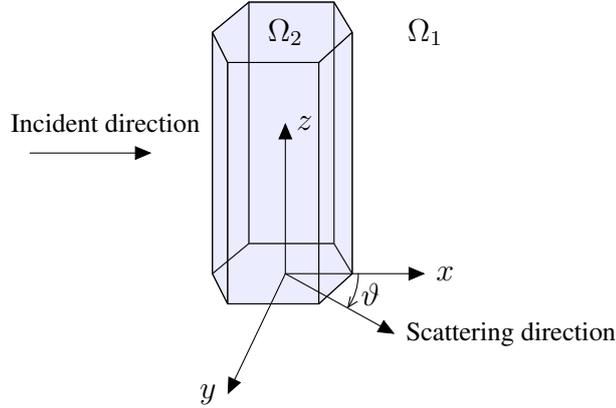
Figure 2.1: Scattering setup

Notation: We employ a combination of the notations used in [26] and [36]. Given a set $\Omega \subset \mathbb{R}^3$, let $\mathbf{L}^2_{loc}(\Omega)$ denote the complex, vector-valued space of locally square integrable functions. We denote by $\mathbf{H}^s_{loc}(\Omega)$, $s \geq 0$ the standard Sobolev spaces with the convention $\mathbf{H}^0 \equiv \mathbf{L}^2$ (for definitions, see, e.g., [103], [1]). The suffix loc is removed when Ω is bounded, i.e., for spaces on Ω_2. We note that for the acoustic case, when the functions under consideration are scalar valued, the boldface letters for function spaces will be replaced by regular type face. Let $\mathbf{D}$ denote a differential operator, then we define for $s \geq 0$:

$$\mathbf{H}^s_{loc}(\mathbf{D}, \Omega) := \{\mathbf{u} \in \mathbf{H}^s_{loc}(\Omega) : \ \mathbf{D}\mathbf{u} \in \mathbf{H}^s_{loc}(\Omega)\}. \tag{2.23}$$

For $s = 0$, we write $\mathbf{H}^0 = \mathbf{H}$ for simplicity. If Ω is bounded, $\mathbf{H}^s_{loc}(\mathbf{D}, \Omega)$ is endowed with the natural graph norm $||\mathbf{u}||^2_{\mathbf{H}^s(\mathbf{D},\Omega)} := ||\mathbf{u}||^2_{\mathbf{H}^s(\Omega)} + ||\mathbf{D}\mathbf{u}||^2_{H^s(\Omega)}$. This defines in particular the Hilbert spaces $\mathbf{H}(\mathbf{curl}, \Omega)$ and $\mathbf{H}(\mathbf{curl}\,\mathbf{curl}, \Omega)$.

Now we may state the *electromagnetic transmission problem*.

Given an incident field $\mathbf{E}^i(\mathbf{x}) = \mathbf{E}_0 e^{ik_1 \mathbf{d}^i \cdot \mathbf{x}}$, where $\mathbf{x} \in \mathbb{R}^3$ and $\mathbf{d}^i$ is a unit direction vector, we wish to determine the fields $\mathbf{E}_1(\mathbf{x}) \in \mathbf{H}_{loc}(\mathbf{curl}, \overline{\Omega}_1)$ and $\mathbf{E}_2(\mathbf{x}) \in \mathbf{H}(\mathbf{curl}, \Omega_2)$ satisfying

$$\nabla \times \nabla \times \mathbf{E}_1 - k_1^2 \mathbf{E}_1 = 0 \quad \text{in } \Omega_1, \tag{2.24}$$

$$\nabla \times \nabla \times \mathbf{E}_2 - k_2^2 \mathbf{E}_2 = 0 \quad \text{in } \Omega_2, \tag{2.25}$$

and the transmission conditions on the boundary Γ:

$$\gamma_D^+ \mathbf{E}_1 = \gamma_D^- \mathbf{E}_2 \quad \text{and} \quad \frac{k_1}{\mu_1} \gamma_N^+ \mathbf{E}_1 = \frac{k_2}{\mu_2} \gamma_N^- \mathbf{E}_2 \quad \text{on } \Gamma, \tag{2.26}$$

22

where $k_j = \omega\sqrt{\mu_j \varepsilon_j}$ for $j = 1, 2$. In addition, the scattered field $\mathbf{E}^s := \mathbf{E}_1 - \mathbf{E}^i$ must satisfy the Silver-Müller radiation condition

$$\hat{\mathbf{x}} \times \nabla \times \mathbf{E}^s(\mathbf{x}) + ik_1\mathbf{E}^s(\mathbf{x}) = o\left(\frac{1}{r}\right) \quad \text{as} \quad r := |\mathbf{x}| \to \infty \tag{2.27}$$

uniformly in all directions $\hat{\mathbf{x}} := \mathbf{x}/r$. The above operators γ_D^+ and γ_D^- denote respectively the exterior and interior Dirichlet trace operators where the Dirichlet trace operator $\gamma_D :$ $\mathbf{H}(\mathbf{curl}; \Omega) \to \mathbf{H}_\times^{-1/2}(\Gamma)$ is the unique bounded and continuous linear mapping for which

$$\gamma_D\mathbf{E} := \mathbf{n} \times \mathbf{E}|_\Gamma,$$

if $\mathbf{E} \in \mathbf{C}^\infty(\overline{\Omega})$ and where $\mathbf{n}$ is the outward normal to Γ. The Sobolev space $\mathbf{H}_\times^{-1/2}(\Gamma)$ is the trace space defined in [26, Definition 1]. The operators γ_N^+ and γ_N^- denote respectively the exterior and interior Neumann trace operators where the Neumann trace operator $\gamma_N :$ $\mathbf{H}(\mathbf{curl}; \Omega) \to \mathbf{H}_\times^{-1/2}(\Gamma)$ is the unique bounded and continuous linear mapping for which

$$\gamma_N\mathbf{E} := \frac{1}{ik}\mathbf{n} \times (\nabla \times \mathbf{E})|_\Gamma,$$

if $\mathbf{E} \in \mathbf{C}^\infty(\overline{\Omega})$. Note that here (for the Neumann trace) we are following [124]. This differs from the definition in [26] in that it includes an extra i factor in the denominator. For more details on trace operators for vector-valued functions on Lipschitz domains, the reader is referred to [25] and [26].

Theorem 2.1 (Existence and uniqueness). *(See [37, Proposition 4.7].)* *Let $k_1, k_2 \in \mathbb{C}$ and $\alpha := \mu_1/\mu_2 \in \mathbb{C}\backslash\{0\}$ be such that either*

$\quad k_1 > 0, \; \text{Im}(\alpha) \leq 0 \text{ and } \text{Im}(\alpha k_2^2) \geq 0,$

or

$\quad \text{Im}(k_1) > 0 \text{ or } k_1 = 0$, *and there exist no $a, b, c, d > 0$ with $-\alpha k_2^2 a - k_1^2 b + \alpha c + d = 0$.* *Then a solution to the electromagnetic transmission problem (2.24)–(2.27) exists and is unique.*

The a, b, c, d are integrals of positive-valued functions. For more details, see [37, Proposition 4.7] and the accompanying proof.

Proof. The proof follows similar arguments to that for the acoustic case shown in [37, Proposition 4.7]. The main difference being that here one must use the integration by parts formula for vector functions $\mathbf{U}$ and $\mathbf{V}$,

$$\int_\Omega (\nabla \times \mathbf{U}) \cdot \mathbf{V}\mathrm{dx} = \int_\Omega \mathbf{U} \cdot \nabla \times \mathbf{V}\mathrm{dx} - \int_{\partial\Omega} (\mathbf{U} \times \mathbf{n}) \cdot \mathbf{V}\mathrm{d}s \tag{2.28}$$

rather than Green's first identity (see, e.g., [36]).

$\square$

Remark 2.1. *The conditions under which the uniqueness theorem holds are met by the examples considered in this thesis. These examples all involve the scattering of EM radiation by an ice crystal surrounded by air. The magnetic permeability of air and ice are both unity, i.e., $\mu_1 = \mu_2 = 1$. The refractive index of air is $n_1 = 1$, and the refractive index of ice is such that $\mathrm{Re}(n_2) > 0$ and $\mathrm{Im}(n_2) \geq 0$. So α, k_1, k_2 satisfy*

$$k_1 > 0, \quad \alpha = 1, \quad \mathrm{Re}(k_2) > 0, \quad \mathrm{Im}(k_2) \geq 0.$$

Hence we have that $k_1 > 0$, $\mathrm{Im}(\alpha) \leq 0$ and $\mathrm{Im}(\alpha k_2^2) \geq 0$. Thus the conditions of Theorem 2.1 are satisfied.

2.4 Electromagnetic boundary integral equations

We briefly summarise the reformulation of the transmission problem as a system of boundary integral equations. The exposition here follows that of [124] which in turn closely follows [26]; the latter gives an in-depth review of the boundary integral formulation for this problem.

The first step in deriving a (direct) boundary integral equation formulation is to write down representation formulae for the fields in each of the two domains, Ω_1 and Ω_2, in terms of their unknown boundary data. To do this we introduce the single-layer potential operator Ψ_{SL} and double-layer potential operator Ψ_{DL}:

$$\Psi_{SL} v(\mathbf{x}) := \mathrm{i}k \int_\Gamma G_k(\mathbf{x}, \mathbf{y}) v(\mathbf{y}) d\Gamma(\mathbf{y})$$
$$- \frac{1}{\mathrm{i}k} \nabla \int_\Gamma G_k(\mathbf{x}, \mathbf{y}) \nabla \cdot v(\mathbf{y}) d\Gamma(\mathbf{y}) \tag{2.29}$$

and

$$\Psi_{DL} v(\mathbf{x}) := \nabla \times \int_\Gamma G_k(\mathbf{x}, \mathbf{y}) v(\mathbf{y}) d\Gamma(\mathbf{y}), \tag{2.30}$$

where G_k is the Green's function of the Helmholtz equation with wavenumber k:

$$G_k(\mathbf{x}, \mathbf{y}) = \frac{\exp(\mathrm{i}k|\mathbf{x} - \mathbf{y}|)}{4\pi|\mathbf{x} - \mathbf{y}|}.$$

Now an integral representation for $\mathbf{E}$ in Ω_1 can be written succinctly as

$$\mathbf{E}(\mathbf{x}) = \Psi_{DL} \gamma_D^+ \mathbf{E}(\mathbf{x}) + \Psi_{SL} \gamma_N^- \mathbf{E}(\mathbf{x}), \quad \mathbf{x} \in \Omega_1, \tag{2.31}$$

where Ψ_{SL} and Ψ_{DL} are as in (2.29) and (2.30), respectively, with $k = k_1$. Similarly, $\mathbf{E}$ in Ω_2 possesses the representation

$$\mathbf{E}(\mathbf{x}) = -\Psi_{DL} \gamma_D^+ \mathbf{E}(\mathbf{x}) - \Psi_{SL} \gamma_N^+ \mathbf{E}(\mathbf{x}), \quad \mathbf{x} \in \Omega_2, \tag{2.32}$$

where Ψ_{SL} and Ψ_{DL} are as in (2.29) and (2.30), respectively, with $k = k_2$. These are the Stratton-Chu representation formulae and are often written in a more explicit form without potential operators. For these expressions the reader is referred to [101] and [111].

Boundary integral operators are constructed by applying the Dirichlet and Neumann traces $\gamma_D^\pm, \gamma_N^\pm$ to the potentials Ψ_{SL} and Ψ_{DL}. We might expect to obtain four different boundary integral operators, however due to the fact that $\nabla \times \Psi_{SL} = k\Psi_{DL}$ and $\nabla \times \Psi_{DL} = k\Psi_{SL}$, we have that

$$\gamma_N^\pm \Psi_{SL} = \gamma_D^\pm \Psi_{DL}, \quad \gamma_N^\pm \Psi_{DL} = \gamma_D^\pm \Psi_{SL}.$$

Hence two different boundary integral operators are sufficient for electromagnetic scattering. We define them as

$$\mathcal{R}_\pm := \gamma_D^\pm \Psi_{SL} = \gamma_N^\pm \Psi_{DL}, \qquad \mathcal{C}_\pm := \gamma_D^\pm \Psi_{DL} = \gamma_N^\pm \Psi_{SL}.$$

Applying the interior and exterior traces to the Stratton-Chu formulae, we arrive at the boundary integral equations [26]

$$\left(-\frac{1}{2}I + \mathcal{C}_-\right)\gamma_D^- \mathbf{E} + \mathcal{R}_- \gamma_N^- \mathbf{E} = 0 \tag{2.33}$$

$$-\mathcal{R}_- \gamma_D^- \mathbf{E} + \left(-\frac{1}{2}I + \mathcal{C}_-\right)\gamma_N^- \mathbf{E} = 0 \tag{2.34}$$

$$\left(\frac{1}{2}I + \mathcal{C}_+\right)\gamma_D^+ \mathbf{E} + \mathcal{R}_+ \gamma_N^+ \mathbf{E} = \gamma_D^+ \mathbf{E}^i \tag{2.35}$$

$$-\mathcal{R}_+ \gamma_D^+ \mathbf{E} + \left(\frac{1}{2}I + \mathcal{C}_+\right)\gamma_N^+ \mathbf{E} = \gamma_N^+ \mathbf{E}^i. \tag{2.36}$$

Writing (2.33) and (2.34) in terms of exterior traces by using the transmission conditions $\gamma_D^+ u = \gamma_D^- u$, $\frac{k_2}{\mu_2}\gamma_N^+ u = \frac{k_1}{\mu_1}\gamma_N^- u$, and writing $\rho = \frac{\mu_2 k_1}{\mu_1 k_2}$, gives

$$\left(-\frac{1}{2}I + \mathcal{R}_-\right)\gamma_D^+ \mathbf{E} + \frac{1}{\rho}\mathcal{R}_- \gamma_N^+ \mathbf{E} = 0 \tag{2.37}$$

$$-\mathcal{R}_- \gamma_D^+ \mathbf{E} + \frac{1}{\rho}\left(-\frac{1}{2}I + \mathcal{C}_-\right)\gamma_N^+ \mathbf{E} = 0. \tag{2.38}$$

We have arrived at the four boundary integral equations (2.37), (2.38), (2.35) and (2.36) for the two unknowns $\gamma_D \mathbf{E}$ and $\gamma_N \mathbf{E}$. There are numerous ways to select two equations or two linear combinations of these; see [101] for a discussion of five of these. Here we choose the simplest, the *combined-field formulation*, which is known to be uniquely solvable [102]. This consists of taking the two combinations

$$(2.35) + (2.37) \quad \text{and} \quad (2.36) + (2.38).$$

25

After dropping the $\pm$ on the traces, we have the system of boundary integral equations to be solved:

$$\begin{pmatrix} \mathcal{C}_+ + \mathcal{C}_- & \mathcal{R}_+ + \frac{1}{\rho}\mathcal{R}_- \\ -\mathcal{R}_+ - \rho\mathcal{R}_- & \mathcal{C}_+ + \mathcal{C}_- \end{pmatrix} \begin{pmatrix} \gamma_D \mathbf{E} \\ \gamma_N \mathbf{E} \end{pmatrix} = \begin{pmatrix} \gamma_D \mathbf{E}^i \\ \gamma_N \mathbf{E}^i \end{pmatrix}. \tag{2.39}$$

Once this system is solved for $\mathbf{E}$ and its normal derivative on Γ, the solution is simply substituted into (2.31) or (2.32) to obtain the solution anywhere in $\mathbb{R}^3$.

2.5 Two-dimensional problems

In this thesis we are mainly concerned with time-harmonic scattering by a two-dimensional penetrable obstacle. We shall demonstrate that when the dependence of the electric and magnetic fields is reduced by one dimension, the electromagnetic scattering problem is equivalent to two scalar (or acoustic) transmission problems, one for each of the electric and magnetic fields. This provides our motivation for studying the acoustic problem in two dimensions.

We begin by assuming that the vectors $\mathbf{E}$ and $\mathbf{H}$ in three dimensional space have no z-dependence, i.e.,
$\mathbf{E} = \mathbf{E}(x, y)$ and $\mathbf{H} = \mathbf{H}(x, y)$, although they may have z-components. It is convenient to introduce the notation

$$\mathbf{E}_\perp := (0, 0, E_z), \tag{2.40}$$

$$\mathbf{E}_\parallel := (E_x, E_y, 0), \tag{2.41}$$

$$\mathbf{H}_\perp := (0, 0, H_z), \tag{2.42}$$

$$\mathbf{H}_\parallel := (H_x, H_y, 0), \tag{2.43}$$

where $\mathbf{E} = (E_x, E_y, E_z)$ and $\mathbf{H} = (H_x, H_y, H_z)$ so that $\mathbf{E} = \mathbf{E}_\perp + \mathbf{E}_\parallel$ and $\mathbf{H} = \mathbf{H}_\perp + \mathbf{H}_\parallel$. We also assume that the scatterer Γ is an infinite prism stretching to $z = \pm\infty$, hence its geometry is independent of the z-direction. The outward pointing normal to $\mathbf{n} = (n_x, n_y, 0)$.

The first two spatial components of (2.10) give us

$$(E_x, E_y) = \frac{\mathrm{i}}{\omega\varepsilon} \left(\frac{\partial}{\partial y}, -\frac{\partial}{\partial x} \right) H_z. \tag{2.44}$$

Hence it is only necessary to solve for $\mathbf{H}_\perp$ (i.e., H_z) in order to find $\mathbf{E}_\parallel$. Similarly, the first two spatial components of (2.11) gives

$$(H_x, H_y) = \frac{\mathrm{i}}{\omega\mu} \left(-\frac{\partial}{\partial y}, \frac{\partial}{\partial x} \right) E_z. \tag{2.45}$$

So knowledge of $\mathbf{E}_\perp$ (i.e., E_z) allows us to determine $\mathbf{H}_\parallel$. That is, in order to determine the entire EM field, we must solve the two scalar Helmholtz equations for E_z and H_z

$$(\nabla^2 + k^2)E_z = 0, \tag{2.46}$$

$$(\nabla^2 + k^2)H_z = 0. \tag{2.47}$$

In order to ascertain the correct boundary conditions for E_z and H_z at the interface Γ, we must first consider the boundary condition (2.20). By the linearity of the curl operator, we have that

$$\mathbf{n} \times [\mathbf{E}_\parallel] + \mathbf{n} \times [\mathbf{E}_\perp] = 0. \tag{2.48}$$

Since $\mathbf{E}_\parallel$ and $\mathbf{E}_\perp$ are orthogonal, we must have that both

$$\mathbf{n} \times [\mathbf{E}_\parallel] = 0 \tag{2.49}$$

and

$$\mathbf{n} \times [\mathbf{E}_\perp] = 0. \tag{2.50}$$

Expanding out (2.49) and making use of the relation (2.44) leads to the boundary condition

$$\left[\frac{1}{\varepsilon} \frac{\partial H_z}{\partial \mathbf{n}} \right] = 0. \tag{2.51}$$

Expanding out (2.50) leads to the boundary condition

$$[E_z] = 0. \tag{2.52}$$

In a similar fashion, the vector boundary condition (2.19) leads to the two scalar boundary conditions

$$\left[\frac{1}{\mu} \frac{\partial E_z}{\partial \mathbf{n}} \right] = 0, \tag{2.53}$$

$$[H_z] = 0. \tag{2.54}$$

To summarise, the 2D electromagnetic problem is equivalent to solving two scalar Helmholtz equations

$$(\Delta + k^2)E_z = 0, \tag{2.55}$$

$$(\Delta + k^2)H_z = 0, \tag{2.56}$$

subject to the boundary conditions

$$[E_z] = 0, \tag{2.57}$$

$$[H_z] = 0, \tag{2.58}$$

$$\left[\frac{1}{\mu} \frac{\partial E_z}{\partial \mathbf{n}} \right] = 0, \tag{2.59}$$

$$\left[\frac{1}{\varepsilon} \frac{\partial H_z}{\partial \mathbf{n}} \right] = 0. \tag{2.60}$$

2.6 The (2D) acoustic transmission problem

Wave scattering problems in which the unknown scattered field is a scalar function arise in acoustical applications. Therefore, we take the time to restate our boundary value problem in a manner in keeping with the acoustics literature. This task is also warranted by the fact that there exist some fundamental differences between two and three-dimensional problems which lead to different integral operators and representation theorems. The most notable differences are that the fundamental solution of the Helmholtz equation is now a Hankel function as opposed to a spherical function, and that the boundary integral equations for the transmission problem now involve four boundary integral operators as opposed to just two in the EM problem.

The *acoustic transmission problem* is stated as follows. Let Ω_2 denote a convex polygon in 2D with boundary Γ and outward pointing normal $\mathbf{n}$, where a convex polygon is a convex subset of $\mathbb{R}^2$ enclosed by intersecting straight-line segments. Let $\Omega_1 := \mathbb{R}^d \backslash \overline{\Omega}_2$ denote the exterior unbounded domain. Given $k_1, k_2, \alpha \in \mathbb{C}$ and $F \in H^{1/2}(\Gamma)$, $G \in H^{-1/2}(\Gamma)$, find $u_1 \in H^1_{loc}(\Omega_1)$ and $u_2 \in H^1(\Omega_2)$ such that

$$(\Delta + k_1^2)u_1 = 0, \quad \text{in } \Omega_1, \tag{2.61}$$

$$(\Delta + k_2^2)u_2 = 0, \quad \text{in } \Omega_2, \tag{2.62}$$

$$\gamma^- u_2 - \gamma^+ u_1 = F, \quad \text{on } \Gamma, \tag{2.63}$$

$$\alpha \partial_n^- u_2 - \partial_n^+ u_1 = G \quad \text{on } \Gamma, \tag{2.64}$$

with u_1 satisfying the *Sommerfeld radiation condition*, i.e.,

$$\frac{\partial u_1}{\partial r}(\mathbf{x}) - \mathrm{i}k_1 u_1(\mathbf{x}) = o(r^{-(d-1)/2}) \tag{2.65}$$

as $r := |\mathbf{x}| \to \infty$, uniformly in $\hat{\mathbf{x}} := \mathbf{x}/r$. The operators γ^+ and γ^- denote respectively the exterior and interior Dirichlet trace operators where the Dirichlet trace operator $\gamma : H^1(\Omega) \to H^{1/2}(\Gamma)$ is the unique bounded linear operator defined by

$$\gamma u := u|_\Gamma, \tag{2.66}$$

if $u \in C^\infty(\overline{\Omega})$. The operators ∂_n^+ and ∂_n^- denote respectively the exterior and interior Neumann trace operators where the Neumann trace operator $\partial_n : H^1(\Omega) \to L^2(\Gamma)$ is the unique bounded linear operator defined by

$$\partial_n u := \nabla u \cdot \mathbf{n}|_\Gamma, \tag{2.67}$$

if $u \in C^\infty(\overline{\Omega})$ and with $\mathbf{n}$ the normal to Γ pointing from Ω_2 into Ω_1. For more information on trace operators for scalar functions on Lipschitz domains the reader is referred to [103].
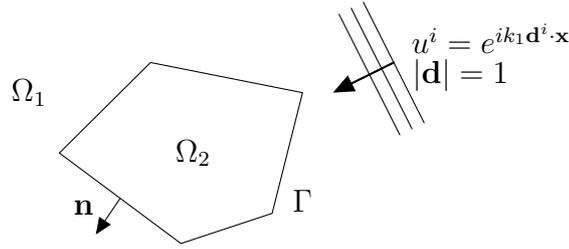
Figure 2.2: Scattering of a plane wave by a 2D polygon.

The *acoustic scattering problem* of interest to us, namely a plane wave $u^i(\mathbf{x}) = \mathrm{e}^{ik_1\mathbf{d}^i\cdot\mathbf{x}}$ incident upon a polygon as in Figure 2.2, can be formulated in this way with the functions F and G in (2.63)–(2.64) being given by

$$F = \gamma^+ u^i \quad \text{and} \quad G = \partial_n^+ u^i. \tag{2.68}$$

The total field is given by

$$u(\mathbf{x}) = \begin{cases} u^i(\mathbf{x}) + u_1(\mathbf{x}), & \mathbf{x} \in \Omega_1. \\ u_2(\mathbf{x}), & \mathbf{x} \in \Omega_2, \end{cases} \tag{2.69}$$

and the transmission conditions (2.63)–(2.64) can be written in terms of the total field as

$$\gamma^+ u = \gamma^- u, \tag{2.70}$$

$$\partial_n^+ u = \alpha \partial_n^- u. \tag{2.71}$$

Sufficient conditions on the constants k_1, k_2, α ensuring unique solvability of (2.61)–(2.65) are provided by the following theorem which we quote from [88, Proposition 2.1 and Corollary 3.4], which follows from results in [37, 133], and also the related result of [99, Corollary 8.5] (note the similarity to Theorem 2.1).

Theorem 2.2 (Existence and uniqueness). *Let $k_1, k_2 \in \mathbb{C}$ and $\alpha \in \mathbb{C}\backslash\{0\}$ be such that either*

$k_1 > 0, \ \mathrm{Im}(\alpha) \leq 0 \ and \ \mathrm{Im}(\alpha k_2^2) \geq 0,$

or

$\mathrm{Im}(k_1) > 0 \ or \ k_1 = 0 \ and \ there \ exist \ no \ a, b, c, d > 0 \ with \ -\alpha k_2^2 a - k_1^2 b + \alpha c + d = 0.$
Then a solution to the problem (2.61)–(2.65) exists and is unique.

Remark 2.2. *The problems considered in this thesis shall be based upon the application to light scattering by ice crystals for which the exterior medium Ω_1 is air which has unit refractive index. Hence $k_1 \in \mathbb{R}$. The interior medium Ω_2 is ice which has a complex*

refractive index $n = n_R + \mathrm{i} n_I$ with $n_R > 0$ and $n_I \geq 0$. So $k_2 := n k_1 \in \mathbb{C}$ with $\mathrm{Re}(\mathrm{k}_2) > 0$ and $\mathrm{Im}(\mathrm{k}_2) \geq 0$. As mentioned earlier, the 2D electromagnetic problem can be decoupled into two 2D acoustic problems - one for the electric field and one for the magnetic field. The difference between these two problems, as can be seen in [52], is the jump α in the solution's normal derivative across Γ.

For the electric problem, $\alpha = 1$, and we can immediately see that the first set of conditions in Theorem 2.2 hold. For the magnetic problem, $\alpha = \sqrt{k_1/k_2}$. For this case, it is simple to show that $\mathrm{Im}(\alpha) \leq$ and $\mathrm{Im}(\alpha \mathrm{k}_2^2) \geq 0$ and hence the first set of conditions in Theorem 2.2 hold.

We highlight here that in this thesis, we shall concentrate on the 2D magnetic problem, i.e., $\alpha = 1$. This is done for simplicity and we leave the verification of the method for $\alpha \neq 1$ to future work. Nevertheless, we derive the integral equations in §2.7 and the geometrical optics formulae in §4 for general $\alpha \in \mathbb{C}$ so that the implementation of the method for $\alpha \neq 1$ should be straightforward.

2.7 Acoustic boundary integral equations

We now reformulate the BVP (2.61)–(2.65), with F and G as given in (2.68), as a system of boundary integral equations via Green's representation theorem. In order to state Green's representation theorem, we first define the single-layer and double-layer potentials for $\phi \in L^1(\Gamma)$ as (for further details see [28, §2.2])

$$S_j \phi(\mathbf{x}) := \int_\Gamma \Phi_j(\mathbf{x}, \mathbf{y}) \phi(\mathbf{y}) \mathrm{d}s(\mathbf{y}), \quad \mathbf{x} \in \mathbb{R}^d \backslash \Gamma, \tag{2.72}$$

$$D_j \phi(\mathbf{x}) := \int_\Gamma \frac{\partial \Phi_j(\mathbf{x}, \mathbf{y})}{\partial n(\mathbf{y}))} \phi(\mathbf{y}) \mathrm{d}s(\mathbf{y}), \quad \mathbf{x} \in \mathbb{R}^d \backslash \Gamma, \tag{2.73}$$

where Φ_j is given by

$$\Phi_j(\mathbf{x}, \mathbf{y}) := \frac{\mathrm{i}}{4} H_0^{(1)}(k_j |\mathbf{x} - \mathbf{y}|)$$

in 2D for $j = 1, 2$.

A form of Green's representation theorem holds for the total field in the interior domain Ω_2 [28, Theorem 2.20]

$$S_2 \partial_n^- u(\mathbf{x}) - D_2 \gamma^- u(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \Omega_1. \\ u(\mathbf{x}), & \mathbf{x} \in \Omega_2. \end{cases} \tag{2.74}$$

Similarly, in the exterior domain Ω_1, for the scattered field we have [28, Theorem 2.21]

$$-S_1 \partial_n^+ u_1(\mathbf{x}) + D_1 \gamma^+ u_1(\mathbf{x}) = \begin{cases} u_1(\mathbf{x}), & \mathbf{x} \in \Omega_1. \\ 0, & \mathbf{x} \in \Omega_2. \end{cases} \tag{2.75}$$

In order to obtain an integral representation for the total field in Ω_1, as opposed to merely the scattered field, we first note that from [28, Theorem 2.20], since u^i is a C^∞ solution of $(\nabla + k_1^2)u^i = 0$ in a neighbourhood of Ω_2,

$$0 = S_1\partial_n^+ u^i(\mathbf{x}) - D_1\gamma^+ u^i(\mathbf{x}) = 0, \quad \mathbf{x} \in \Omega_1. \tag{2.76}$$

Adding (2.76) to (2.75) and then adding $u^i(\mathbf{x})$ to both sides, we find that

$$u(\mathbf{x}) = u^i(\mathbf{x}) - S_1\partial_n^+ u(\mathbf{x}) + D_1\gamma^+ u(\mathbf{x}), \ \mathbf{x} \in \Omega_1. \tag{2.77}$$

Combining (2.74) and (2.77) we have a representation for u in the entire domain $\mathbb{R}^2$, namely

$$u(\mathbf{x}) = \begin{cases} u^i(\mathbf{x}) - S_1\partial_n^+ u(\mathbf{x}) + D_1\gamma^+ u(\mathbf{x}), & \mathbf{x} \in \Omega_1. \\ S_2\partial_n^- u(\mathbf{x}) - D_2\gamma^- u(\mathbf{x}), & \mathbf{x} \in \Omega_2. \end{cases} \tag{2.78}$$

Now to obtain integral equations, we take the Dirichlet and Neumann traces of u onto Γ. We require the standard jump relations for layer potentials which we quote from [28, p. 115]. We have on $H^{-1/2}(\Gamma)$

$$\gamma^\pm S_j = \mathcal{S}_j, \tag{2.79}$$

$$\partial_n^\pm S_j = \mp\frac{1}{2}I + \mathcal{D}'_j, \tag{2.80}$$

and, on $H^{1/2}(\Gamma)$,

$$\gamma^\pm D_j = \pm\frac{1}{2}I + \mathcal{D}_j, \tag{2.81}$$

and

$$\partial_n^\pm D_j\psi = \mathcal{H}_j\psi. \tag{2.82}$$

Here I is the identity operator and $\mathcal{S}_j, \mathcal{D}_j, \mathcal{D}'_j, \mathcal{H}_j$, for $j = 1, 2$ are, respectively, the single-layer, double-layer, adjoint double-layer and hypersingular integral operators defined for $\phi \in L^2(\Gamma)$ as (see [28, § 2.3] for technical details)

$$\mathcal{S}_j\phi(\mathbf{x}) := \int_\Gamma \Phi_j(\mathbf{x}, \mathbf{y})\phi(\mathbf{y})\mathrm{d}s(\mathbf{y}), \tag{2.83}$$

$$\mathcal{D}_j\phi(\mathbf{x}) := \int_\Gamma \frac{\partial\Phi_j(\mathbf{x}, \mathbf{y})}{\partial n(\mathbf{y})}\phi(\mathbf{y})\mathrm{d}s(\mathbf{y}), \tag{2.84}$$

$$\mathcal{D}'_j\phi(\mathbf{x}) := \int_\Gamma \frac{\partial\Phi_j(\mathbf{x}, \mathbf{y})}{\partial n(\mathbf{x})}\phi(\mathbf{y})\mathrm{d}s(\mathbf{y}), \tag{2.85}$$

$$\mathcal{H}_j\phi(\mathbf{x}) := \frac{\partial}{\partial n(\mathbf{x})}\int_\Gamma \frac{\partial\Phi_j(\mathbf{x}, \mathbf{y})}{\partial n(\mathbf{y})}\phi(\mathbf{y})\mathrm{d}s(\mathbf{y}). \tag{2.86}$$

Applying both the Dirichlet traces γ^+ and γ^- to (2.78) and employing the jump relations (2.79) and (2.81) yields the boundary integral equations

$$\left(\frac{1}{2}I - \mathcal{D}_1\right)\gamma^+ u + \mathcal{S}_1 \partial_n^+ u = \gamma^+ u^i, \tag{2.87}$$

$$\left(\frac{1}{2}I + \mathcal{D}_2\right)\gamma^- u - \mathcal{S}_2 \partial_n^- u = 0, \tag{2.88}$$

where we have also dropped the explicit dependence on $\mathbf{x}$ for brevity of presentation. Now applying both the Neumann traces ∂_n^+ and ∂_n^- to (2.78) and employing the jump relations (2.80) and (2.82) yields the boundary integral equations

$$\left(\frac{1}{2}I + \mathcal{D}_1'\right)\partial_n^+ u - \mathcal{H}_1 \gamma^+ u = \partial_n^+ u^i, \tag{2.89}$$

$$\left(\frac{1}{2}I - \mathcal{D}_2'\right)\partial_n^- u + \mathcal{H}_2 \gamma^- u = 0. \tag{2.90}$$

We shall write all the equations henceforth in terms of positive traces. To this end, we employ the transmission conditions (2.70)–(2.71) in equations (2.88) and (2.90). We also drop the trace notation, for the sake of simplifying the presentation, by writing $u(\mathbf{x}) = \gamma^+ u(\mathbf{x})$ and $\partial u(\mathbf{x})/\partial \mathbf{n} = \partial_n^+ u(\mathbf{x})$ when $\mathbf{x} \in \Gamma$. Thus, equations (2.87)–(2.90) are written as

$$\left(\frac{1}{2}I - \mathcal{D}_1\right)u + \mathcal{S}_1 \frac{\partial u}{\partial n} = u^i, \tag{2.91}$$

$$\left(\frac{1}{2}I + \mathcal{D}_2\right)u - \frac{1}{\alpha}\mathcal{S}_2 \frac{\partial u}{\partial \mathbf{n}} = 0, \tag{2.92}$$

$$\left(\frac{1}{2}I + \mathcal{D}_1'\right)\frac{\partial u}{\partial \mathbf{n}} - \mathcal{H}_1 u = \frac{\partial u^i}{\partial \mathbf{n}}, \tag{2.93}$$

$$\left(\frac{1}{2}I - \mathcal{D}_2'\right)\frac{\partial u}{\partial \mathbf{n}} + \alpha \mathcal{H}_2 u = 0. \tag{2.94}$$

These are four boundary integral equations in two unknowns u and $\partial u/\partial n$. Which equations, or linear combination of equations, should we solve? One simple choice would be to solve only (2.91) and (2.92). But this pair is known to not always be uniquely solvable [106]. The choice we make follows that of [83] and [82] (where the smooth scatterer case was considered), and is motivated by the desire to obtain a pair of second kind integral equations with weakly singular kernels by making the singularites in the hypersingular operators partially cancel. We do this by taking the combinations (2.91)+(2.92) and α(2.93)+(2.94) to give

$$\frac{1}{2}(I + \alpha I)u + (\alpha \mathcal{D}_2 - \mathcal{D}_1)u + (\mathcal{S}_1 - \mathcal{S}_2)\frac{\partial u}{\partial n} = u^i, \tag{2.95}$$

$$\frac{1}{2}(I + \alpha I)\frac{\partial u}{\partial n} + (\alpha \mathcal{D}_1' - \mathcal{D}_2')\frac{\partial u}{\partial n} + \alpha(\mathcal{H}_2 - \mathcal{H}_1)u = \alpha \frac{\partial u^i}{\partial n}. \tag{2.96}$$

This cancellation of the strong singularities in $\mathcal{H}_1$ and $\mathcal{H}_2$ is attractive from both an analytic and a computational viewpoint.

To recap, the solution of the transmission problem (2.61)–(2.65), with F and G given by (2.68), satisfies Green's representation formula

$$u(\mathbf{x}) = \begin{cases} u^i(\mathbf{x}) - S_1 \frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}) + D_1 u(\mathbf{x}), & \mathbf{x} \in \Omega_1. \\ S_2 \frac{1}{\alpha} \frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}) - D_2 u(\mathbf{x}), & \mathbf{x} \in \Omega_2, \end{cases} \tag{2.97}$$

and the boundary integral equation

$$\mathcal{A}\mathbf{v} = \mathbf{f}, \tag{2.98}$$

where

$$\mathbf{v} = \begin{pmatrix} u \\ \frac{\partial u}{\partial \mathbf{n}} \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} u^i \\ \alpha \frac{\partial u^i}{\partial \mathbf{n}} \end{pmatrix},$$

and

$$\mathcal{A} = \begin{pmatrix} \frac{1}{2}(1+\alpha)I + (\alpha \mathcal{D}_2 - \mathcal{D}_1) & \mathcal{S}_1 - \mathcal{S}_2 \\ \alpha(\mathcal{H}_2 - \mathcal{H}_1) & \frac{1}{2}(1+\alpha)I + (\alpha \mathcal{D}_1' - \mathcal{D}_2') \end{pmatrix}. \tag{2.99}$$

The boundary integral operator $\mathcal{A}$ can be viewed as a mapping

$$\mathcal{A} : H^{1/2}(\Gamma) \times H^{-1/2}(\Gamma) \to H^{1/2}(\Gamma) \times H^{-1/2}(\Gamma),$$

$$\mathcal{A} : H^1(\Gamma) \times L^2(\Gamma) \to H^1(\Gamma) \times L^2(\Gamma),$$

or

$$\mathcal{A} : L^2(\Gamma) \times L^2(\Gamma) \to L^2(\Gamma) \times L^2(\Gamma).$$

It is desirable here to work in the final context since this will be the most simple for the Galerkin implementation we shall describe later in the thesis. This choice is reasonable since the data $\mathbf{f}$ are always in $L^2(\Gamma) \times L^2(\Gamma)$ for the problems we consider.

We shall now prove that for the relevant k_1, k_2, α (as described in Remark 2.2), $\mathcal{A}$ is invertible on $L^2(\Gamma) \times L^2(\Gamma)$. This will complete the proof of the following theorem.

Theorem 2.3. *For $k_1 > 0$, $\mathrm{Re}(\mathrm{k}_2) > 0$, $\mathrm{Im}(\mathrm{k}_2) \geq 0$, and either*

(i) $\alpha = 1$, or

(ii) $\alpha = \sqrt{k_1/k_2} = 1/\sqrt{n}$,

the solution of the acoustic scattering problem exists, is unique, and is given by Green's representation theorem (2.97), where u and $\partial u/\partial \mathbf{n}$ satisfy the boundary integral equation (2.98). Futhermore, the boundary integral operator $\mathcal{A}$ is invertible on $L^2(\Gamma) \times L^2(\Gamma)$.

Proof. It only remains to prove the final statement of the theorem, i.e. that $\mathcal{A}$ is invertible on $L^2(\Gamma) \times L^2(\Gamma)$. This is shown using a modification of the Fredholm-type argument employed in [133] where an indirect formulation for the three-dimensional acoustic transmission problem for Lipschitz scatterers is considered.

We first sketch an outline of the proof before providing the details. We begin by decomposing the integral operator $\mathcal{A}$ as $\mathcal{A} = \mathcal{J} + \mathcal{K}$ where $\mathcal{J}$ is shown to be an invertible operator on $L^2 \times L^2$ (thus trivially a Fredholm operator of index zero), and $\mathcal{K}$ is shown to be a compact operator on $L^2 \times L^2$ via perturbations from the Laplace case. Hence $\mathcal{A} = \mathcal{J} + \mathcal{K}$ is Fredholm of index zero (by [103, Theorem 2.26]) and so the Fredholm Alternative ([103, Theorem 2.27(i)]) is available to us. Thus, we endeavour to show that $\mathcal{A}$ is injective on $L^2 \times L^2$ and then conclude, by the Fredholm Alternative, that $\mathcal{A}$ is invertible on $L^2 \times L^2$.

Now to commence with the proof of the invertibility of $\mathcal{A}$. First let us introduce the $k = 0$ versions of the single-layer, double-layer and adjoint double-layer operators and denote them as $\mathcal{S}_0$, $\mathcal{D}_0$ and $\mathcal{D}_0'$, respectively. Then notice that

$$\frac{1}{2}(1+\alpha)I + (\alpha\mathcal{D}_2 - \mathcal{D}_1) = \frac{1}{2}(1+\alpha)I - (1-\alpha)\mathcal{D}_0 + (\mathcal{D}_0 - \mathcal{D}_1) + \alpha(\mathcal{D}_2 - \mathcal{D}_0) \quad (2.100)$$

and

$$\frac{1}{2}(1+\alpha)I + (\alpha\mathcal{D}_1' - \mathcal{D}_2') = \frac{1}{2}(1+\alpha)I - (1-\alpha)\mathcal{D}_0' + (\mathcal{D}_0' - \mathcal{D}_2') + \alpha(\mathcal{D}_1' - \mathcal{D}_0'). \quad (2.101)$$

Therefore we may write

$$\mathcal{A} = \mathcal{J} + \mathcal{K} \quad (2.102)$$

where

$$\mathcal{J} = \begin{pmatrix} \frac{1}{2}(1+\alpha)I - (1-\alpha)\mathcal{D}_0 & 0 \\ 0 & \frac{1}{2}(1+\alpha)I - (1-\alpha)\mathcal{D}_0' \end{pmatrix} \quad (2.103)$$

and

$$\mathcal{K} = \begin{pmatrix} (\mathcal{D}_0 - \mathcal{D}_1) + \alpha(\mathcal{D}_2 - \mathcal{D}_0) & \mathcal{S}_1 - \mathcal{S}_2 \\ \alpha(\mathcal{H}_2 - \mathcal{H}_1) & (\mathcal{D}_0' - \mathcal{D}_2') + \alpha(\mathcal{D}_1' - \mathcal{D}_0') \end{pmatrix}. \quad (2.104)$$

For $\alpha = 1$, $\mathcal{J}$ is clearly invertible. In order to show $\mathcal{J}$ is invertible for $\alpha \neq 1$, we shall use the following result quoted from [46, Theorem 2.2(i)].

Theorem 2.4. *Assume Ω is a bounded Lipschitz convex domain in $\mathbb{R}^n$, $n \geq 2$. Then the spectral radius of $\mathcal{D}_0'$ on $L^2(\partial\Omega)$ equals $\frac{1}{2}$.*

An immediate consequence of Theorem 2.4 is that $\lambda I - \mathcal{D}_0'$ is invertible for $|\lambda| > \frac{1}{2}$. This statement also holds for $\lambda I - \mathcal{D}_0$. A further consequence of Theorem 2.4 is the following lemma, which we quote from [133, Lemma 3.2].

Lemma 2.1. *Suppose $\lambda I - \mathcal{D}_0$ is invertible in $L^2(\Gamma)$, then $\lambda I - \mathcal{D}_0$ is invertible in $H^1(\Gamma)$.*

From the fact that $\lambda I - \mathcal{D}_0$ is invertible on $L^2(\Gamma)$ (a consequence of Theorem 2.4), we can deduce that the operator $\mathcal{J}$ is invertible if

$$\left| \frac{1+\alpha}{1-\alpha} \right| > 1. \tag{2.105}$$

The values of α that we are interested in are those such that $\alpha = n^{-1/2}$ where $\mathrm{Re}(n) > 0$ and $\mathrm{Im}(n) \geq 0$. So we may write $\alpha = |n|^{-1/2}e^{i\theta}$ where $-\pi/4 \leq \theta \leq 0$ and hence we have that $\mathrm{Re}(\alpha) > 0$. Since $\mathrm{Re}(\alpha) > 0$, the condition (2.105) holds and $\mathcal{J}$ is invertible.

Now we wish to show that $\mathcal{K}$ is compact. In order to do this we require the following known results.

(i) $\mathcal{S}_k - \mathcal{S}_0 : L^2(\Gamma) \to L^2(\Gamma)$ is compact [28, p. 122],

(ii) $\mathcal{D}_k - \mathcal{D}_0 : L^2(\Gamma) \to L^2(\Gamma)$ is compact [133, p.1466]
(and hence so is $\mathcal{D}'_k - \mathcal{D}'_0 : L^2(\Gamma) \to L^2(\Gamma))$,

(iii) $\mathcal{H}_k - \mathcal{H}_0 : L^2(\Gamma) \to L^2(\Gamma)$ is compact [133, Lemma 6.2],

(iv) $\mathcal{D}_k - \mathcal{D}_0 : L^2(\Gamma) \to H^1(\Gamma)$ is bounded [133, p. 1466].

Using these results, we can immediately see that each operator entry in $\mathcal{K}$ is compact, and hence $\mathcal{K}$ is compact on $L^2(\Gamma) \times L^2(\Gamma)$. Therefore $\mathcal{A}$ is a Fredholm operator of index zero and thus, by the Fredholm alternative (see, e.g., [103, Theorem 2.27(i)]), the system of boundary integral equations (2.98) has a unique solution if the homogeneous system $(\mathbf{f} = 0)$ has only the trivial solution. In other words, we must show that $\mathcal{A}$ is injective.

To prove the injectivity of $\mathcal{A}$, we begin by supposing that

$$\mathbf{v} = \begin{pmatrix} \phi \\ \psi \end{pmatrix} \in L^2(\Gamma) \times L^2(\Gamma) \tag{2.106}$$

solves the system (2.98) with $u^i = 0$, i.e. $\mathcal{A}\mathbf{v} = \mathbf{0}$, and we endeavour to prove that $\mathbf{v} = \mathbf{0}$. First let us note that ϕ possesses more regularity than we have assumed; in fact we have that $\phi \in H^1(\Gamma)$. This can be seen by expanding out $\mathcal{A}\mathbf{v} = 0$, i.e. $(\mathcal{J} + \mathcal{K})\mathbf{v} = 0$ and considering the first row. This gives us

$$\left(\frac{1}{2}(1+\alpha)I - (1-\alpha)\mathcal{D}_0 \right)\phi = (\mathcal{D}_1 - \mathcal{D}_0)\phi - \alpha(\mathcal{D}_2 - \mathcal{D}_0)\phi - (\mathcal{S}_1 - \mathcal{S}_2)\psi.$$

Since $(\mathcal{D}_1 - \mathcal{D}_0)$ and $(\mathcal{D}_2 - \mathcal{D}_0)$ map $L^2(\Gamma)$ to $H^1(\Gamma)$ (see result (iv) on p.35), and $\mathcal{S}_k$ maps $L^2(\Gamma)$ to $H^1(\Gamma)$ also [133, p.1466], we have that

$$\left(\frac{1}{2}(1+\alpha)I - (1-\alpha)\mathcal{D}_0 \right)\phi \in H^1(\Gamma).$$

Then by Lemma 2.1 the operator $\left(\frac{1}{2}(1 + \alpha)I - (1 - \alpha)\mathcal{D}_0\right)$ is invertible in $H^1(\Gamma)$, showing that $\phi \in H^1(\Gamma)$. This will be necessary later for showing that the functions to be designed are in the required function spaces in order to satisfy the BVP.

Now consider the functions v_1 and v_2 defined by

$$v_1 := \alpha D_2 \phi - S_2 \psi \quad \text{in } \Omega_1, \tag{2.107}$$

$$v_2 := D_1 \phi - S_1 \psi \quad \text{in } \Omega_2, \tag{2.108}$$

where we note that the interior and exterior wavenumbers are switched and we wish to show that these functions satisfy the homogeneous version of the BVP defined by (2.61)–(2.65) and (2.70)–(2.71). Applying the trace operators and using the jump relations (2.79)–(2.82), we get

$$\gamma^+ v_1 = \alpha \left(\frac{1}{2}I + \mathcal{D}_2\right)\phi - \mathcal{S}_2\phi, \tag{2.109}$$

$$\gamma^- v_2 = \left(-\frac{1}{2}I + \mathcal{D}_1\right)\phi - \mathcal{S}_1\psi, \tag{2.110}$$

$$\partial_n^+ v_1 = \alpha \mathcal{H}_2 \phi - \left(-\frac{1}{2}I + \mathcal{D}_2'\right)\psi, \tag{2.111}$$

$$\partial_n^- v_2 = \mathcal{H}_1 \phi - \left(\frac{1}{2}I + \mathcal{D}_1'\right)\psi. \tag{2.112}$$

Then

$$\gamma^+ v_1 = \gamma^- v_2 \quad \text{and} \quad \partial_n^+ v_1 = \alpha \partial_n^- v_2.$$

Since, as was shown earlier, $(\phi, \psi) \in H^1(\Gamma) \times L^2(\Gamma)$, as can be shown via the well-known mapping properties of the single- and double-layer potentials (see, e.g., [28, Theorem 2.15]) that

$$S_k \psi \in H^1_{loc}(\mathbb{R}^2) \quad \text{and} \quad D_k \phi \in H^1_{loc}(\Omega_k)$$

for $k = 1, 2$. Thus $v_1 \in H^1_{loc}(\Omega_1)$ and $v_2 \in H^1(\Omega_2)$. Hence we see that v_1, v_2 also satisfy the homogeneous version of the BVP defined by (2.61)–(2.65) and (2.70)–(2.71), and so by Theorem 2.2 (uniqueness) we have that $v_1 \equiv u_2 \equiv 0$.

Next consider the functions u_1 and u_2 defined by

$$u_1 := D_1 \phi - S_1 \psi \quad \text{in } \Omega_1, \tag{2.113}$$

$$u_2 := -D_2 \phi - \frac{1}{\alpha} S_2 \psi \quad \text{in } \Omega_2. \tag{2.114}$$

Note here that (2.113) and (2.114) are identical to u as defined in (2.97) for $u^i \equiv 0$. Applying the trace operators and using the jump relations gives

$$\gamma^+ u_1 = \alpha\left(\frac{1}{2}I + \mathcal{D}_1\right)\phi - \mathcal{S}_1\phi, \tag{2.115}$$

$$\gamma^- u_2 = \left(\frac{1}{2}I - \mathcal{D}_2\right)\phi + \frac{1}{\alpha}\mathcal{S}_2\psi, \tag{2.116}$$

$$\partial_n^+ u_1 = \alpha\mathcal{H}_1\phi - \left(\frac{1}{2}I - \mathcal{D}_1'\right)\psi, \tag{2.117}$$

$$\partial_n^- u_2 = -\mathcal{H}_2\phi + \frac{1}{\alpha}\left(\frac{1}{2}I + \mathcal{D}_2'\right)\psi. \tag{2.118}$$

Then we see that

$$\begin{aligned}
\gamma^+ u_1 - \gamma^- u_2 &= (\mathcal{D}_1 + \mathcal{D}_2)\phi - (\mathcal{S}_1 + \frac{1}{\alpha}\mathcal{S}_2)\psi \\
&= \frac{1}{\alpha}\gamma^+ v_1 + \gamma^- v_2 \\
&= 0 \quad \text{since } v_1 \equiv v_2 \equiv 0,
\end{aligned}$$

and

$$\begin{aligned}
\partial_n^+ u_1 - \partial_n^- u_2 &= (\mathcal{H}_1 + \alpha\mathcal{H}_2)\phi - (\mathcal{D}_1' + \mathcal{D}_2')\psi \\
&= \partial_n^+ v_1 + \partial_n^- v_2 \\
&= 0 \quad \text{since } v_1 \equiv v_2 \equiv 0.
\end{aligned}$$

Further, by the same use of [28, Theorem 2.15] as above, we have that $u_1 \in H^1_{loc}(\Omega_1)$ and $u_2 \in H^1(\Omega_2)$. Hence we see that u_1, u_2 also satisfy the homogeneous BVP. Therefore by uniqueness, $u_1 \equiv u_2 \equiv 0$.

Now we have all the pieces in place for the final step of the proof. Recall that we are supposing that

$$\mathcal{A}\begin{pmatrix}\phi \\ \psi\end{pmatrix} = \mathbf{0}.$$

Expanding this out, we have

$$\frac{1}{2}(1+\alpha)\phi = (\mathcal{S}_2 - \mathcal{S}_1)\psi - (\alpha\mathcal{D}_2 - \mathcal{D}_1)\phi, \tag{2.119}$$

$$\frac{1}{2}(1+\alpha)\psi = \alpha(\mathcal{H}_1 - \mathcal{H}_2)\phi + (\mathcal{D}_2' - \alpha\mathcal{D}_1')\psi, \tag{2.120}$$

37

which after some manipulation yield

$$
\begin{aligned}
\phi &= (\mathcal{S}_2 - \mathcal{S}_1)\psi - (\alpha\mathcal{D}_2 - \mathcal{D}_1)\phi - \frac{\alpha}{2}\phi + \frac{\phi}{2} \\
&= \underbrace{\mathcal{S}_2\psi - \alpha\left(\frac{1}{2}I + \mathcal{D}_2\right)\phi}_{-\gamma^+ v_1} + \underbrace{\left(\frac{1}{2}I + \mathcal{D}_1\right)\phi - \mathcal{S}_1\psi}_{\gamma^+ u_1} \\
&= 0, \quad \text{since } u_1 \equiv v_1 \equiv 0,
\end{aligned}
$$

and

$$
\begin{aligned}
\alpha\psi &= \alpha(\mathcal{H}_1 - \mathcal{H}_2)\phi + (\mathcal{D}_2' - \alpha\mathcal{D}_1')\psi + \frac{\alpha}{2}\psi - \frac{1}{2}\psi \\
&= \underbrace{-\alpha\mathcal{H}_2\phi + \left(-\frac{1}{2}I + \mathcal{D}_2'\right)\psi}_{-\partial_n^+ v_1} + \underbrace{\alpha\mathcal{H}_1\phi - \alpha\left(-\frac{1}{2}I + \mathcal{D}_1'\right)\psi}_{\alpha\partial_n^+ u_1} \\
&= 0,
\end{aligned}
$$

hence $\psi = 0$ since we have assumed $\alpha \neq 0$. Therefore $\mathbf{v} = \mathbf{0}$ and the BIE system is uniquely solvable. $\qquad\square$

2.8   Scattering properties

When solving scattering problems via (direct) boundary integral equations, the solution we obtain is the field and its normal derivative only on the surface of the scatterer. However, the boundary data is often of little practical interest. Rather it is the field in the scattering domain away from the scattering surface which is the desired output. This field can be calculated by substituting the boundary data into the Stratton-Chu formulae (2.31)–(2.32) in the EM case, or Green's representation formulae (2.78) in the acoustic (and 2D EM) case. However, when the point of observation is "far" from the scatterer, i.e., $k_1 r \gg 1$, where $|r| := |\mathbf{x}|$, the asymptotic versions of the aforementioned formulae are employed. We shall state these far-field formulae in this section. We also detail the *amplitude scattering matrix* and its calculation which is relevant for the EM scattering problem. This matrix describes how the far-field amplitudes of the scattered radiation's different polarisation states are related to those of the incident wave. We also discuss briefly the scattering phase function and scattering cross section which are quantities of interest in EM scattering simulations. For more detail on these and other scattering properties, the reader is referred to [18] or [142].

2.8.1 Far-field pattern

Recall the Stratton-Chu formula for $\mathbf{E}^s \in \Omega_1$,

$$\mathbf{E}^s(\mathbf{x}) = -\Psi_{DL}\gamma_D^+\mathbf{E}(\mathbf{x}) - \Psi_{SL}\gamma_N^+\mathbf{E}(\mathbf{x}), \quad \mathbf{x} \in \Omega_1. \tag{2.121}$$

When $k_1 r \gg 1$, this may be written in the asymptotic form [111]

$$\mathbf{E}^s(\mathbf{x}) \sim -\frac{e^{ik_1 r}}{ik_1 r}\mathbf{F}(\hat{\mathbf{d}}^s), \quad \text{as } kr \to \infty, \tag{2.122}$$

where $\hat{\mathbf{d}}^s := \mathbf{x}/|\mathbf{x}| \in \mathbb{S}^2$ is the unit scattering direction, with the *far-field pattern* being given by

$$\mathbf{F} = \Psi_{DL}^F(\gamma_D^+\mathbf{E}) + \Psi_{SL}^F(\gamma_N^+\mathbf{E}),$$

where the far-field versions of the single- and double-layer potentials are defined as

$$[\Psi_{DL}^F v](\hat{\mathbf{d}}^s) = \frac{k_1^2}{4\pi}\int_\Gamma e^{-ik(\hat{\mathbf{d}}^s \cdot \mathbf{y})}\left(\hat{\mathbf{d}}^s \times v(\mathbf{y})\right) d\Gamma(\mathbf{y}) \tag{2.123}$$

$$[\Psi_{SL}^F v](\hat{\mathbf{d}}^s) = \frac{k_1^2}{4\pi}\int_\Gamma e^{-ik(\hat{\mathbf{d}}^s \cdot \mathbf{y})}v(\mathbf{y}) d\Gamma(\mathbf{y})$$
$$- \frac{k_1^2}{4\pi}\int_\Gamma e^{-ik(\hat{\mathbf{d}}^s \cdot \mathbf{y})}v(\mathbf{y}) \cdot \hat{\mathbf{d}}^s d\Gamma(\mathbf{y}). \tag{2.124}$$

It is customary to define the scattering direction $\hat{\mathbf{d}}^s$ in terms of the two angles ϑ and ϕ. The angle ϕ dictates the scattering plane and is a rotation about the x-axis, and ϑ is the angle from the forward scattering direction which describes the scattering direction within that plane. In the 3D scattering examples presented in §3, the forward scattering direction is always taken to be the positive x-direction. Therefore, the scattering direction $\hat{\mathbf{d}}^s$ is written

$$\hat{\mathbf{d}}^s = (\cos(\vartheta), \cos(\phi)\sin(\vartheta), \sin(\phi)\cos(\vartheta)). \tag{2.125}$$

In the examples shown in §3, we shall consider, for simplicity, the plane $\phi = 0$, i.e., the (x, y)-plane, hence

$$\hat{\mathbf{d}}^s = (\cos(\vartheta), \sin(\vartheta), 0)$$

for $0 < \vartheta < 2\pi$.

For the acoustic case, recall Green's representation formula for $u^s \in \Omega_1$,

$$u^s(\mathbf{x}) = -S_1\partial_n^+ u(\mathbf{x}) + D_1\gamma^+ u(\mathbf{x}), \ \mathbf{x} \in \Omega_1. \tag{2.126}$$

The so-called *far-field* is obtained by using the asymptotic behaviour of the Hankel functions $H_0^{(1)}$ and $H_1^{(1)}$ for large argument [39] within Green's representation formula for the scattered field (2.75). This gives

$$u^{ff}(\mathbf{x}) \sim \frac{e^{i\pi/4}}{2\sqrt{2\pi}}\frac{e^{ik_1 r}}{\sqrt{k_1 r}}F(\hat{\mathbf{d}}^s), \ \text{as } r := |\mathbf{x}| \to \infty, \tag{2.127}$$

39

where $\hat{\mathbf{d}}^s := \mathbf{x}/|\mathbf{x}| \in \mathbb{S}^1$, the unit circle, with the *far-field pattern* F being given by

$$F(\hat{\mathbf{d}}^s) = -\int_\Gamma \mathrm{e}^{-\mathrm{i}k_1\hat{\mathbf{d}}^s\cdot\mathbf{y}} \left(\mathrm{i}k_1(\hat{\mathbf{d}}^s \cdot \mathbf{n}(\mathbf{y}))u(\mathbf{y}) + \frac{\partial u}{\partial \mathbf{n}}(\mathbf{y}) \right) \mathrm{d}s(\mathbf{y}). \tag{2.128}$$

Here the scattering direction $\hat{\mathbf{d}}^s$ is only dependent on $\vartheta \in [0, 2\pi)$ and is

$$\hat{\mathbf{d}}^s = (\cos(\vartheta), \sin(\vartheta)).$$

2.8.2 Amplitude scattering matrix

Before describing the amplitude scattering matrix, it is necessary to define the reference and scattering planes. We choose the reference plane to be defined by the incident wave direction $\mathbf{d}^i$ and two perpendicular unit vectors $\hat{\mathbf{e}}_\parallel^i$ and $\hat{\mathbf{e}}_\perp^i$ such that $\hat{\mathbf{e}}_\perp^i \times \hat{\mathbf{e}}_\parallel^i = \mathbf{d}^i$. Here, for simplicity (as mentioned in the previous section), we take $\mathbf{d}^i = \hat{\mathbf{x}}$, and so $\hat{\mathbf{e}}_\parallel^i = \hat{\mathbf{y}}$ and $\hat{\mathbf{e}}_\perp^i = \hat{\mathbf{z}}$. The scattering plane is the plane containing both $\mathbf{d}^i$ and the scattering direction $\mathbf{d}^s$, and we define vectors $\hat{\mathbf{e}}_\perp^s$ and $\hat{\mathbf{e}}_\parallel^s$ such that $\hat{\mathbf{e}}_\perp^s \times \hat{\mathbf{e}}_\parallel^s = \mathbf{d}^s$. The angle between the two planes is denoted by ϕ. We shall here take $\phi = 0$ for simplicity, and the resulting reference and scattering planes are both the (x, y)-plane with vectors $\hat{\mathbf{e}}_\parallel^i$, $\hat{\mathbf{e}}_\perp^i$, $\hat{\mathbf{e}}_\parallel^s$, $\hat{\mathbf{e}}_\perp^s$ depicted in Figure 2.3.
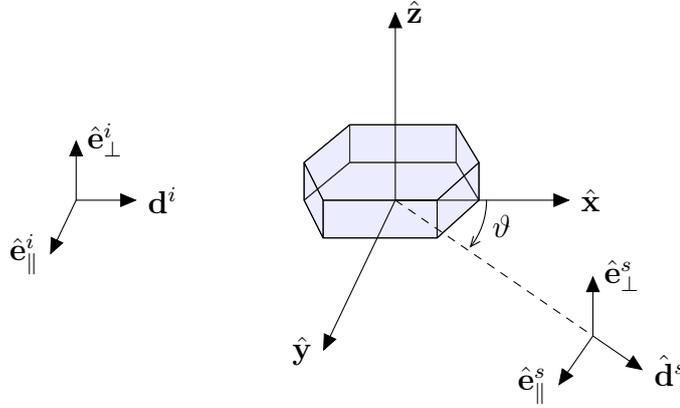
Figure 2.3: Reference and scattering planes.

Now the incident and scattered fields may be decomposed into parallel and perpendicular components in the following way

$$\mathbf{E}^i(\mathbf{x}) = \left(E_\parallel^i \hat{\mathbf{e}}_\parallel^i + E_\perp^i \hat{\mathbf{e}}_\perp^i \right) \mathrm{e}^{\mathrm{i}k_1 x}, \tag{2.129}$$

$$\mathbf{E}^s(\mathbf{x}) = \left(E_\parallel^s \hat{\mathbf{e}}_\parallel^s + E_\perp^s \hat{\mathbf{e}}_\perp^s \right) \mathrm{e}^{\mathrm{i}k_1 r}. \tag{2.130}$$

We note that this decomposition for $\mathbf{E}^s$ is possible since the scattered field is approximately transverse in the far-field, that is, $\hat{\mathbf{d}}^s \cdot \mathbf{F} \approx 0$ (see, e.g., [73]). $E_\parallel^i$ and $E_\perp^i$ are scalar, complex-valued electric amplitudes of the parallel and perpendicular components of the incident electric field vector, and $E_\parallel^s$ and $E_\perp^s$ are their counterparts for the scattered electric field vector. These quantities are related via the *amplitude scattering matrix* in the following way:

$$\begin{pmatrix} E_{\parallel s}^s \\ E_{\perp s}^s \end{pmatrix} = \frac{\mathrm{e}^{\mathrm{i}k_1(r-x)}}{-\mathrm{i}k_1 r} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} E_{\parallel i}^i \\ E_{\perp i}^i \end{pmatrix} \cdot \tag{2.131}$$

It may be readily observed that if the incident field is polarised in either of the $\hat{\mathbf{y}}$- or $\hat{\mathbf{z}}$-directions, then the amplitude scattering matrix only has two non-zero entries. A simple way to calculate all entries A_{ij}, is to solve the scattering twice, once with each of the aforementioned polarisations of the incident wave. For example, we may take the two incident fields

$$\mathbf{E}_1^i = (1 \cdot \hat{\mathbf{e}}_{\perp i} + 0 \cdot \hat{\mathbf{e}}_{\parallel i}) e^{\mathrm{i}k_1 x} \tag{2.132}$$

and

$$\mathbf{E}_2^i = (1 \cdot \hat{\mathbf{e}}_{\parallel i} + 0 \cdot \hat{\mathbf{e}}_{\perp i}) e^{\mathrm{i}k_1 x}. \tag{2.133}$$

From the solution to the scattering problem with incident field $\mathbf{E}_1^i$, one may calculate A_{12} and A_{22}, and from the solution to the scattering problem with incident field $\mathbf{E}_2^i$, one may calculate A_{12} and A_{21}.

From the amplitude scattering matrix, the entries of the so-called *scattering matrix* can be calculated. The scattering matrix relates the Stokes parameters of the incident and scattered fields and the reader is referred to an EM scattering monograph such as [18] or [105] for details of these quantities. We shall only state the first entry of the scattering matrix which provides a measure of the scattered intensity and is analogous to the square of the far-field pattern (2.128) in the acoustic case. This entry is often called the *phase function* and is given by

$$S_{11} = \frac{1}{2}(|A_{11}|^2 + |A_{22}|^2 + |A_{21}|^2 + |A_{12}|^2). \tag{2.134}$$

This entry is often normalised in the following manner:

$$P_{11} = \frac{4\pi S_{11}}{k_1^2 C_{sca}}, \tag{2.135}$$

where C_{sca} is the *scattering cross section* defined as

$$C_{sca} = \int_0^{2\pi} \int_0^\pi \frac{S_{11}}{k_1^2} \sin\vartheta \mathrm{d}\vartheta \mathrm{d}\phi, \tag{2.136}$$

that is, the scattering cross section is the integral of the far-field amplitude over all scattering directions.

41

Chapter 3

Numerical methods

In §1.2.1 a brief review of popular numerical methods for wave scattering problems was given. Each method has its own particular advantages and disadvantages, but, in their conventional form, they all have in common a computational cost which grows with increasing size parameter χ. In this chapter we shall describe in detail one particular numerical method - the boundary element method. The #DOF for the conventional BEM, which employs piecewise polynomial basis functions, scales with χ as $\mathcal{O}(\chi^{d-1})$ where $d = 2, 3$ is the dimension of the scattering problem. However, it is applicable to any particle geometry, is relatively simple to implement, and has excellent and well-understood convergence properties.

The outline of this chapter is as follows. In §3.1 we describe the boundary element method in general terms. Such a description is pertinent also to the HNA approach which lies within the BEM framework. §3.2 details the implementation of a conventional 2D Galerkin hp-BEM for scattering by convex polygons. This includes in §3.2.2 a discussion of the evaluation of the singular integral operator differences arising in our 2D boundary integral equation choice (2.99). In particular, we see explicitly the cancellation of the strong singularities in the hypersingular operators. In §3.2.3 we detail efficient quadrature methods for evaluating the arising singular integrals. All of these implementation techniques are relevant also for the HNA BEM of §6. §3.2.4 demonstrates the convergence properties of the 2D Galerkin hp-BEM which shall be employed in the following chapters to generate the necessary reference solutions.

Finally, §3.3 briefly describes the use of a 3D open-source BEM code, BEM++ [124], to solve the 3D EM transmission problem. The purpose of this demonstration is to illustrate the extremely demanding computational cost of a state-of-the-art implementation of a BEM for a simple 3D scattering problem. In particular, we demonstrate that, to achieve approximately 1% accuracy in the far-field pattern for scattering by a hexagonal ice column,

the limit in terms of χ when using a quad-core PC with 16.4 Gigabytes of RAM is $\chi = 15$. This provides motivation for the development of HNA methods.

3.1 Boundary element method

We briefly outline the main ideas of the BEM in two of its main forms - the collocation and Galerkin methods - following the exposition given in [5, Chapter 3].

Consider the integral equation

$$\mathcal{K}u = b, \tag{3.1}$$

where $\mathcal{K}$ is an integral operator mapping a Banach space $\mathcal{X}$ into itself. For the context we work in in this thesis, $\mathcal{X} = L^2(\Gamma) \times L^2(\Gamma)$.

We begin by choosing a finite dimensional subspace $\mathcal{X}_N \subset \mathcal{X}$ where N denotes the dimension of $\mathcal{X}_N$. Let $\mathcal{X}_N$ have a basis $\{\phi_1, \ldots, \phi_N\}$. Then we seek a function $v_N \in \mathcal{X}_N$ which may be written as

$$v_N(\mathbf{x}) = \sum_{i=1}^{N} c_i \phi_i(\mathbf{x}), \quad \mathbf{x} \in \Gamma. \tag{3.2}$$

The function representation (3.2) is substituted into the integral equation (3.1) and we may define the *residual* $r_N(\mathbf{x})$ as

$$r_N(\mathbf{x}) = \sum_{i=1}^{N} c_i \mathcal{K} \phi_i(\mathbf{x}) - b(\mathbf{x}). \tag{3.3}$$

The coefficients $\{c_1, \ldots, c_N\}$ are chosen by forcing the residual $r_N(\mathbf{x})$ to be approximately zero in some sense. The two different "senses" we describe define the collocation and Galerkin methods.

3.1.1 Collocation method

In the collocation method, one chooses N distinct *collocation points* $\mathbf{y}_1, \ldots, \mathbf{y}_N$ on the boundary Γ, and then enforces the residual to be zero at these points. This leads to the determination of $\{c_1, \ldots, c_N\}$ via the solution of the linear system

$$\sum_{j=1}^{N} c_j \mathcal{K} \phi_j(\mathbf{y}_i) = b(\mathbf{y}_i), \quad i = 1, \ldots, N. \tag{3.4}$$

Or, written in matrix form, we are required to solve

$$\mathbf{K}\mathbf{v} = \mathbf{b}, \tag{3.5}$$

where the entries in the matrix K are defined by

$$\mathrm{K}_{i,j} = \mathcal{K}\phi_j(\mathbf{y}_i) \tag{3.6}$$

and $\mathbf{v}$ and $\mathbf{b}$ are column vectors with their entries defined as

$$\mathbf{v}_i = c_i \quad \text{and} \quad \mathbf{b}_i = b(\mathbf{y}_i).$$

The collocation method is extremely popular owing to its conceptual simplicity. Also, the integrals to be performed in the left-hand side of (3.4) are all of dimension $d - 1$ for $d = 2, 3$, and so the method is relatively straightforward to implement. However we note that the stability of the method can depend strongly on the choice of collocation points $\mathbf{y}_i$ (see [115], although we note that there the author is employing an HNA approximation space).

3.1.2 Galerkin's method

Another approach to determine the coefficients $\{c_1, \ldots, c_N\}$ is Galerkin's method. Let $\langle \cdot, \cdot \rangle$ denote the usual inner product on $L^2(\Gamma)$, defined by

$$\langle f, g \rangle := \int_\Gamma f(\mathbf{x})\overline{g}(\mathbf{x})\mathrm{d}\Gamma(\mathbf{x}).$$

Then in Galerkin's method we require that the residual $r_N(\mathbf{x})$ satisfy

$$\langle r_N, \phi_i \rangle = 0, \quad i = 1, \ldots, N. \tag{3.7}$$

This leads to the determination of the coefficients c_i via the solution of the linear system

$$\sum_{j=1}^{N} c_j \langle \mathcal{K}\phi_j, \phi_i \rangle = \langle b, \phi_i \rangle, \quad i = 1, \ldots, N. \tag{3.8}$$

Or, written in matrix form, we are required to solve

$$\mathrm{K}\mathbf{v} = \mathbf{b}, \tag{3.9}$$

where the entries in the matrix K are defined by

$$\mathrm{K}_{i,j} = \langle \mathcal{K}\phi_j, \phi_i \rangle \tag{3.10}$$

and $\mathbf{v}$ and $\mathbf{b}$ are column vectors with their entries defined as

$$\mathbf{v}_i = c_i \quad \text{and} \quad \mathbf{b}_i = \langle b, \phi_i \rangle.$$

To implement the Galerkin method we now are required to evaluate integrals of dimension $2(d-1)$ for $d = 2, 3$, as opposed to $d-1$ for the collocation method, hence making the Galerkin method more costly (although only in terms of CPU time and not in terms of memory). However, this method possesses many advantages over the collocation method. In particular, for many scattering problems it is possible to prove that the Galerkin solution is quasi-optimal, i.e., it produces an approximation that is equivalent to the best approximation, up to a constant (although such a proof is not available for the transmission problem). Also, one does not have to worry about the different possible choices of the collocation points with the potential for ill-conditioned systems if they are poorly chosen. For the latter reason, we choose to employ the Galerkin method in this thesis. However, as will be noted in §6, if the stability issues associated with the collocation points could be resolved, a collocation based implementation of the HNA method described in this thesis could offer significant savings in terms of computation and implementation time. This is because the oscillatory integrals arising would be of lower dimension and hence easier to evaluate using oscillatory quadrature techniques.

3.2   A Galerkin hp-BEM for the 2D transmission problem

In this section we describe the application of the Galerkin BEM to the 2D transmission problem in a "conventional" way. By conventional, we mean that we use piecewise polynomials as the basis functions which populate the approximation space. The method described in this section shall be used to generate reference solutions for later comparison with the HNA BEM. It also has many features in common with the HNA BEM to be proposed later in the thesis so we take the time to describe its implementation in detail here.

Recall that the integral equation we wish to solve may be written as

$$\begin{pmatrix} \frac{1}{2}(1+\alpha)I + (\alpha \mathcal{D}_2 - \mathcal{D}_1) & \mathcal{S}_1 - \mathcal{S}_2 \\ \alpha(\mathcal{H}_2 - \mathcal{H}_1) & \frac{1}{2}(1+\alpha)I + (\alpha \mathcal{D}_1' - \mathcal{D}_2') \end{pmatrix} \begin{pmatrix} u \\ \frac{\partial u}{\partial \mathbf{n}} \end{pmatrix} = \begin{pmatrix} u^i \\ \alpha \frac{\partial u^i}{\partial \mathbf{n}} \end{pmatrix}. \qquad (3.11)$$

In §3.2.1 we choose a finite dimensional space of piecewise polynomials $\{\phi_i\}_{i=1}^{N} \subset L^2(\Gamma)$ so that the two unknowns $u, \partial u/\partial \mathbf{n}$ may be approximated by v, w where

$$u \approx v := \sum_{i=1}^{N} V_i \phi_i \quad \text{and} \quad \frac{\partial u}{\partial \mathbf{n}} \approx w := \sum_{i=1}^{N} W_i \phi_i. \qquad (3.12)$$

Applying the Galerkin method to the equation (3.11) with the representations of the unknowns as in (3.12), yields the matrix system to be solved:

$$A \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{c} \end{pmatrix}, \qquad (3.13)$$

where the entries in matrix A are defined by

$$
A_{i,j} = \begin{cases} \frac{1}{2}(1+\alpha)\langle \phi_j, \phi_i \rangle + \langle (\alpha \mathcal{D}_2 - \mathcal{D}_1)\phi_j, \phi_i \rangle, & i = 1, \ldots, N; \ j = 1, \ldots, N, \\ \langle (\mathcal{S}_1 - \mathcal{S}_2)\phi_j, \phi_i \rangle, & i = 1, \ldots, N; \ j = N+1, \ldots, 2N, \\ \langle \alpha(\mathcal{H}_2 - \mathcal{H}_1)\phi_j, \phi_i \rangle, & i = N+1, \ldots, 2N; \ j = 1, \ldots, N; \\ \frac{1}{2}(1+\alpha)\langle \phi_j, \phi_i \rangle + \langle (\alpha \mathcal{D}_1' - \mathcal{D}_2')\phi_j, \phi_i \rangle, & i = 1, \ldots, N; \ j = 1, \ldots, N, \end{cases}
\tag{3.14}
$$

and **v**, **w**, **b**, **c** are column vectors with their entries defined as

$$
\mathbf{v}_i = V_i, \ \mathbf{w}_i = W_i, \ \mathbf{b}_i = \langle u^i, \phi_i \rangle, \ \mathbf{c}_i = \alpha \left\langle \frac{\partial u^i}{\partial \mathbf{n}}, \phi_i \right\rangle.
\tag{3.15}
$$

In §3.2.1 we choose a suitable approximation space $\{\phi_i\}_{i=1}^N \subset L^2(\Gamma)$ for our scattering problem. §3.2.2 and §3.2.3 then discuss the implementation of the Galerkin approximation (3.13). In particular, §3.2.2 is concerned with the accurate numerical evaluation of the kernels in (3.14). The differences of singular operators can lead to large cancellation errors if not handled correctly. §3.2.3 discusses the 1D and 2D integrals that must be evaluated. This includes an exposition of some efficient singular integral quadrature techniques which, although standard, are rarely presented in detail for specific BIE calculations. §3.2.4 presents numerical results demonstrating the excellent accuracy and convergence properties of this Galerkin approximation.

3.2.1 Approximation space

We build our basis functions from a suitable set of orthogonal polynomials owing to their desirable interpolation properties. Here, in view of our L^2 setting, we choose the Legendre polynomials which are defined on the interval $[-1, 1]$ by the three-term recurrence

$$
L_0(x) = 1,
\tag{3.16}
$$

$$
L_1(x) = x,
\tag{3.17}
$$

$$
L_{j+1}(x) = \frac{2j+1}{j+1}x L_j(x) - \frac{j}{j+1}L_{j-1}(x), \quad j \geq 1.
\tag{3.18}
$$

The Legendre polynomials form an orthogonal basis for $L^2([-1, 1])$, specifically, we have

$$
\int_{-1}^{1} L_j(x) L_k(x)\mathrm{d}x = \begin{cases} 0, & j \neq k \\ \frac{2}{2j+1}, & j = k. \end{cases}
$$

On a general interval $[a, b]$, we work with

$$
L_n \left(\frac{2(x-a)}{b-a} - 1 \right),
\tag{3.19}
$$

where $x \in [a, b]$. When we take the inner product of (3.19) with itself, we get, after a change of variables,

$$\int_a^b \left(L_n \left(\frac{2(x-a)}{b-a} - 1 \right) \right)^2 \mathrm{d}x = \frac{b-a}{2} \int_{-1}^1 (L_n(s))^2 \mathrm{d}s \tag{3.20}$$

$$= \frac{b-a}{2} \frac{2}{2n+1}. \tag{3.21}$$

So, in order to have an orthonormal basis for $L^2([a, b])$, we multiply the polynomials (3.19) by $\sqrt{\frac{2}{b-a}} \sqrt{\frac{2n+1}{2}}$.

In order to describe the mesh we use, it is useful to first define some notation pertaining to the geometry of the polygon. Label the corners of the polygon going anti-clockwise by $\mathbf{P}_n, n = 1, \ldots, N_s$, where N_s is the number of sides. Set $\mathbf{P}_{N_s+1} := \mathbf{P}_1$, and, for $n = 1, \ldots, N_s$, denote by Γ_n the side between corners $\mathbf{P}_n$ and $\mathbf{P}_{n+1}$. The point $\mathbf{x} \in \Gamma$, whose arc length measured anti-clockwise around Γ from $\mathbf{P}_1$ is s, is represented parametrically by

$$\mathbf{x}(s) = \mathbf{P}_n + (s - \tilde{L}_{n-1}) \left(\frac{\mathbf{P}_{n+1} - \mathbf{P}_n}{L_n} \right), \quad \text{for } s \in [\tilde{L}_{n-1}, \tilde{L}_n], \ n = 1, \ldots, N_s, \tag{3.22}$$

where $L_n = |\mathbf{P}_{n+1} - \mathbf{P}_n|$ is the length of side Γ_n, and $\tilde{L}_n = \sum_{m=1}^n L_m, \ n = 1, \ldots, N_s$ denotes the arc-length distance from $\mathbf{P}_1$ to $\mathbf{P}_{n+1}$. We set $\tilde{L}_0 = 0$ and denote the total length of Γ by $L := \tilde{L}_{N_s}$.

Now consider the nth side of the polygon, i.e., Γ_n. The construction of the mesh on Γ_n (for each $n = 1, \ldots, N_s$) is performed in two stages. First, we construct a uniform mesh with $N_e = \lceil \frac{L_n}{\lambda_1} \rceil$ elements, where $\lambda_1 = 2\pi/k_1$ is the wavelength of the incident wave. That is, we have at least two elements per wavelength in the mesh. Let us denote this uniform mesh as $\mathcal{U}_n$. Its meshpoints x_i are defined by

$$x_i := \frac{iL_n}{N_e}, \quad i = 0, 1, \ldots, N_e. \tag{3.23}$$

This mesh ensures that we have a minimum number of two mesh elements per wavelength, however is does not account for potentially low-regularity behaviour of the Cauchy data at the corners of the polygon. We shall make a short digression to discuss this behaviour in order to justify our mesh design.

The behaviour of $(u, \partial u/\partial \mathbf{n})$ at each corner is dependent upon the refractive index of the scatterer, the corner angle, and the complex parameter α (from the transmission condition (2.71)). Some information about this behaviour is presented in [37] however it is not particularly explicit, nor is there any dependence upon the refractive index leading one to discern that their analysis is incomplete. Particularly, one should expect a dependence

on the imaginary part of the refractive index since in the limit of high absorption (as the imaginary part of the refractive index goes to infinity), it can be seen that the transmission problem becomes an impenetrable scattering problem with Dirichlet boundary conditions. For this problem it is known (see, e.g., [63]) that the boundary data $\partial u / \partial \mathbf{n}$ behaves like

$$\left| \frac{\partial u}{\partial \mathbf{n}} \right| \leq C k_1 |k_1 s|^{-\delta}, \quad \text{as } s \to \infty,$$

where s is the distance from the corner of interest, C is a constant, and $\delta = 1 - \pi/\theta$, where θ is the interior corner angle. For such a problem, meshes which are geometrically graded towards the corner are often employed to capture this singular behaviour. For low and zero absorption problems, we do not anticipate such singular behaviour. Nevertheless, in what follows, we shall construct our meshes to account for this high-absorption "worst-cast" scenario.

More precisely, in the construction of our meshes, we replace the two elements nearest the corner points (i.e., those within a wavelength of the corners) with meshes geometrically graded towards $\mathbf{P}_n$ and $\mathbf{P}_{n+1}$. Given $A > 0$, denote by $\mathcal{G}^+_{N_l}(\tilde{L}_n, \tilde{L}_n + A)$ the geometric mesh with N_l layers graded towards $\mathbf{P}_n$, whose meshpoints x_i are defined by

$$x_0 := \tilde{L}_n, \quad x_i := \tilde{L}_n + \sigma^{N_l - i} A, \quad i = 1, 2, \ldots, N_l, \tag{3.24}$$

where $0 < \sigma < 1$ is a grading parameter. Similarly, denote by $\mathcal{G}^-_{N_l}(\tilde{L}_{n+1} - A, \tilde{L}_{n+1})$ the geometric mesh with N_l layers graded towards $\mathbf{P}_{n+1}$, whose meshpoints x_i are defined by

$$x_0 := \tilde{L}_{n+1} - A, \quad x_i := \tilde{L}_{n+1} - \sigma^{N_l - i} A, \quad i = 1, 2, \ldots, N_l. \tag{3.25}$$

A smaller value of the grading parameter represents a more severe grading.

The mesh we use on each side Γ_n is the union of the the three meshes defined above. Denote the final mesh on Γ_n as $\mathcal{M}_n$, then we have that

$$\mathcal{M}_n = \mathcal{U}_n(\tilde{L}_n, \tilde{L}_{n+1}) \cup \mathcal{G}^+_{N_l}(\tilde{L}_n, \tilde{L}_n + L_n/N_e) \cup \mathcal{G}^-_{N_l}(\tilde{L}_{n+1} - L_n/N_e, \tilde{L}_{n+1}), \tag{3.26}$$

where the number of layers N is yet to be specified. An example of this mesh with $N = 3$, and for $k_1 = 3$ and $L_n = 2\pi$ is shown in Figure 3.1.
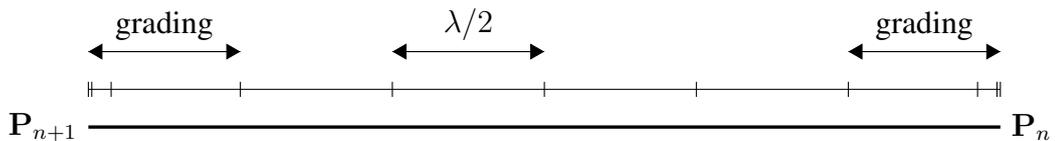
Figure 3.1: Mesh with $N_l = 3$ on a side with $L_n = 2\pi$ and for $k_1 = 3$.

Now we define the space of piecewise polynomials housed on the mesh $\mathcal{M}_n$. We do so in stages, as we did with the meshes. Beginning with the graded meshes; consider the mesh graded towards $\mathbf{P}_n$. Given a vector $\mathbf{p} \in (\mathbb{N}_0)^n$, we let $P_{\mathbf{p},n}(0, A)$ denote the space of piecewise polynomials on the mesh $\mathcal{G}_{n_l}(0, A)$ with the degree vector $\mathbf{p}$, i.e.,

$$\mathcal{P}_{\mathbf{p},n}(0, A) := \big\{ \rho : [0, A] \to \mathbb{C} : \rho|_{(x_{i-1}, x_i)} \text{ is a polynomial of}$$
$$\text{degree less than or equal to } (\mathbf{p})_i, \ i = 1, \ldots, n \big\} .$$

For reasons of efficiency and conditioning it is common to decrease the order of the approximating polynomials towards the singularity. Specifically, in the method proposed here we use a "linear slope" degree vector $\mathbf{p}$ with

$$(\mathbf{p})_i := \begin{cases} p - \left\lfloor \frac{(n+1-i)}{n} p \right\rfloor, & 1 \le i \le n - 1, \\ p, & i = n, \end{cases}$$

where the integer $p \ge 0$ is the highest polynomial degree on the mesh. For the mesh graded towards $\mathbf{P}_{n+1}$, the setup is entirely similar to the above except in this case everything must be flipped so that the grading and slope in $\mathbf{p}$ is toward the corner $\mathbf{P}_{n+1}$.

On the uniform portion of the mesh, each element accommodates polynomials up to and including degree p. This leads to a total number of degrees of freedom in the approximation space given by

$$M := \sum_{j=1}^{N_s} (p+1) \left(\left\lceil \frac{k_1 L_j}{2\pi} \right\rceil + p \right) . \tag{3.27}$$

Now that the approximation space for our conventional hp-BEM has been established, we shall discuss some of the implementational details of the method. In particular, we consider the numerical evaluation of the kernels of the integral operators for the transmission problem. Then we present effective methods for evaluating the integrals necessary for computing the mass matrix.

3.2.2 Kernel evaluation

The BIE formulation (2.95)-(2.96) was chosen since it contains the difference $\mathcal{H}_1 - \mathcal{H}_2$, meaning that for the strong singularities in the hypersingular operators cancel with each other, leaving behind a weakly (logarithmically) singular kernel which may be integrated efficiently using the quadrature techniques presented in the following section. However, close to the singularities of the operators (i.e., when $\mathbf{x} - \mathbf{y}$ is small), large cancellation occurs when performing the subtraction $\mathcal{H}_1 - \mathcal{H}_2$ which may lead to inaccurate kernel evaluations due to numerical rounding error. Near these singularities it is better to use a Taylor series approximation of the difference $\mathcal{H}_1 - \mathcal{H}_2$ directly rather than compute $\mathcal{H}_1$ and

$\mathcal{H}_2$ separately and subtract one from the other. Here we shall present this series approximation for $\mathcal{H}_1 - \mathcal{H}_2$ which is obtained from the well-known series expansions of Hankel functions [39]. This also serves to illustrate the aforementioned cancellation of the strong singularities in the hypersingular operators.

We are interested in the accurate evaluation of the difference

$$(\mathcal{H}_1 - \mathcal{H}_2)\psi(\mathbf{x}) = \frac{\partial}{\partial \mathbf{n}(\mathbf{x})} \int_{\Gamma(\mathbf{y})} \left(\frac{\partial \Phi_1(\mathbf{x}, \mathbf{y})}{\partial \mathbf{n}(\mathbf{y})} - \frac{\partial \Phi_2(\mathbf{x}, \mathbf{y})}{\partial \mathbf{n}(\mathbf{y})} \right) \psi(\mathbf{y}) \mathrm{d}s(\mathbf{y}). \qquad (3.28)$$

It can be shown that

$$\begin{aligned}
\frac{\partial^2 \Phi_k}{\partial \mathbf{n}(\mathbf{x}) \partial \mathbf{n}(\mathbf{y})} = &\frac{\mathrm{i}}{4} \frac{k H_1^{(1)}(kR)}{R} \mathbf{n}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{y}) \\
&+ \frac{\mathrm{i}}{4} \frac{k^2 H_2^{(1)}(kR)}{R^2} ((\mathbf{y} - \mathbf{x}) \cdot \mathbf{n}(\mathbf{x}))((\mathbf{x} - \mathbf{y}) \cdot \mathbf{n}(\mathbf{y})).
\end{aligned} \qquad (3.29)$$

So we see that the difference (3.28) contains the following two terms (multiplied by other factors), namely

$$\Delta_{Ha} = \frac{k_1 H_1^{(1)}(k_1 R)}{R} - \frac{k_2 H_1^{(1)}(k_2 R)}{R}, \qquad (3.30)$$

and

$$\Delta_{Hb} = \frac{k_1^2 H_2^{(1)}(k_1 R)}{R^2} - \frac{k_2^2 H_2^{(1)}(k_2 R)}{R^2}, \qquad (3.31)$$

where $R = |\mathbf{x} - \mathbf{y}|$. Recall $n = k_2/k_1$ is the refractive index of Ω_2, and let us define

$$\rho = k_1 R$$

so that Δ_{Ha} and Δ_{Hb} may be written as

$$\Delta_{Ha} = \frac{k_1^2}{\rho} \left(H_1^{(1)}(\rho) - n H_1^{(1)}(n\rho) \right).$$

and

$$\Delta_{Hb} = \frac{k_1^4}{\rho^2} \left(H_2^{(1)}(\rho) - n^2 H_2^{(1)}(n\rho) \right).$$

We can observe from (3.29) that the Δ_{Hb} term is multiplied by zero when $\mathbf{x}$ and $\mathbf{y}$ are on the same side of the polygon since then $(\mathbf{x} - \mathbf{y}) \cdot \mathbf{n} = 0$. However, when the $\mathbf{x}$ and $\mathbf{y}$ are on different sides but are close to each other at a corner, we need to consider the behaviour of the kernel difference for small R.

Consider the Δ_{Ha} term first. Using the definition $H_\nu^{(1)}(z) := J_\nu(z) + \mathrm{i} Y_\nu(z)$ and the series expansions for $J_\nu(z)$ and $Y_\nu(z)$ for small z as given in [39] with $\nu = 0$, we have that,

for $z < 1$,

$$H_1^{(1)}(z) = -\frac{2\mathrm{i}}{\pi}\frac{1}{z} - \frac{\mathrm{i}}{\pi}\left(\ln 2 - \gamma + 1 - \frac{\pi}{2}\right)z$$
$$+ \frac{\mathrm{i}}{\pi}z\ln z + \frac{\mathrm{i}}{8\pi}\left(\ln 2 - \gamma + \frac{5}{4} + \frac{\mathrm{i}\pi}{2}\right)z^3 \qquad (3.32)$$
$$- \frac{\mathrm{i}}{8\pi}z^3\ln z + \mathcal{O}(z^5).$$

as $z \to 0$, where γ is Euler's constant. Hence, when $\rho \ll 1$, Δ_{Ha} possesses the expansion

$$\Delta_{Ha} \sim -k_1^2\frac{\mathrm{i}}{\pi}\left[\left(\ln 2 - \gamma + 1 - \frac{\pi}{2}\right)(1 - n^2) + n^2\ln n\right]$$
$$+ k_1^2\frac{\mathrm{i}}{\pi}(1 - n^2)\ln\rho + k_1^2\frac{\mathrm{i}}{8\pi}\left[\left(\ln 2 - \gamma + \frac{5}{4} + \frac{\mathrm{i}\pi}{2}\right)(1 - n^4) + n^4\ln n\right]\rho^2$$
$$- k_1^2\frac{\mathrm{i}}{8\pi}(1 - n^4)\rho^2\ln\rho + \mathcal{O}(\rho^4),$$
$$(3.33)$$

which we observe has a logarithmic singularity at $\rho = 0$.

Now consider the Δ_{Hb} term. The function $H_2^{(1)}(z)$ possesses the following series expansion for $z < 1$:

$$H_2^{(1)}(z) = -\frac{4\mathrm{i}}{\pi}\frac{1}{z^2} - \frac{\mathrm{i}}{\pi} - \frac{\mathrm{i}}{4\pi}\left(\ln 2 - \gamma + \frac{3}{4} + \frac{\pi\mathrm{i}}{2}\right)z^2 + \frac{\mathrm{i}}{4\pi}z^2\ln z + \mathcal{O}(z^3) \qquad (3.34)$$

Hence the difference Δ_{Hb} has the expansion

$$\Delta_{Hb} \sim -\frac{\mathrm{i}k_1^4}{4\pi}\left[\left(\ln 2 - \gamma + \frac{3}{4} + \frac{\pi\mathrm{i}}{2}\right)(1 - n^4) + n^4\ln n\right] + \frac{\mathrm{i}k_1^4}{4\pi}(1 - n^4)\ln\rho + \mathcal{O}(\rho)$$
$$(3.35)$$

for $\rho < 1$. It can be seen that the two strong (non-integrable) singularities from the individual hypersingular potentials have cancelled to leave a singularity of the form $x^2\log(x)$.

It is recommended to employ the Taylor expansions (3.33) and (3.35) when $\rho \ll 1$ to avoid round-off errors becoming pervasive. Here we recommend these expansions are used when $\rho = 10^{-4}$ which leads to a relative error of better than 10^{-8} in the evaluation of (3.29).

3.2.3 Quadrature

One of the main drawbacks of the boundary element method is the difficulty in its implementation compared to that of a finite element method. In particular, we are required to approximate integrals numerically, many of which have a singular kernel. Such integrals cannot be approximated efficiently using standard quadrature rules such as Gauss-Legendre, midpoint- or Simpson's rule, and require something more sophisticated. One

approach would be to subdivide the domain of integration according to a geometric grading towards the singularity and then apply a standard quadrature rule on each of the subdivisions. Indeed, this approach is the simplest and most robust method for general integrable singularities. However, if we know the nature of the singularity, it is possible to employ a more efficient approach. We have chosen our BIE formulation specifically so that the kernels all possess singularities of a weak (logarithmic) nature. Thus we may use a generalised Gaussian quadrature (see, e.g., [68]) which is tailor-made for such singularities. We shall detail the application of this approach, as well as the aforementioned geometric grading approach, herein.

In the Galerkin formulation, we must evaluate 2D integrals of the form

$$I = \int_{\Gamma_i} \int_{\Gamma_j} k(\mathbf{x}, \mathbf{y})\phi_j(\mathbf{y})\overline{\phi_i(\mathbf{x})}\mathrm{d}\mathbf{y}\mathrm{d}\mathbf{x}, \tag{3.36}$$

where $k(\mathbf{x}, \mathbf{y})$ is a kernel with a logarithmic singularity at $\mathbf{x} = \mathbf{y}$, ϕ_i and ϕ_j are basis functions with respective supports Γ_i and Γ_j. The integral (3.36) is the general form of those in (3.14). It is helpful for later to rewrite (3.36) in parametric form as

$$I = \int_0^1 \int_a^b p(s)q(t)f(x - y)\mathrm{d}y\mathrm{d}x, \tag{3.37}$$

where the function $f(x - y)$ has a logarithmic singularity at $x = y$, $p(x) \in \mathbb{C}^\infty$ and $q(x) \in \mathbb{C}^\infty$ are typically polynomials for our purposes, and a, b are such that $a \geq 0$ and $b \geq a + 1$. We note that such a rewriting of (3.36) is always possible after appropriate scalings and potentially switching the order of integration.

There are also 1D integrals which arise when computing the right-hand side of (3.13). These take the form

$$I_{rhs} = \int_\Gamma u^i(\mathbf{x})\phi(\mathbf{x})\mathrm{d}\mathbf{x}. \tag{3.38}$$

These 1D integrals are non-singular and may be easily and efficiently evaluated using, for example, Gaussian quadrature. In fact, when u^i is a plane wave, these integrals can be evaluated analytically. Owing to their simplicity, we shall discuss them no further here.

The 2D integrals of the form (3.37) may be categorised in four cases corresponding to different integration domains defined by a and b. These cases are

(i) $a - 1 > \varepsilon$ (ε small): $[0, 1]$ and $[a, b]$ are well-separated,

(ii) $[a, b] = [0, 1]$: singularity on diagonal $x = y$,

(iii) $a = 1$: corner singularity at $x = 1$, $a = 1$,

(iv) $0 < a - 1 < \varepsilon$ (ε small): near singularity.

In case (i), the integral (3.38) may be efficiently evaluated by taking tensor product of 2D Gaussian quadrature rules. Cases (ii), (iii) and (iv) require a more sophisticated treatment which we shall discuss in stages. First we shall present two quadrature rules for 1D integrals with logarithmic singularities or near-singularities, and then show how they may be used via tensor products to efficiently tackle the 2D integrals in cases (ii), (iii) and (iv).

3.2.3.1 Singular one-dimensional integrals

Consider the integral

$$I := \int_0^1 p(x) f(x) \mathrm{d}x \tag{3.39}$$

where f has a logarithmic singularity at $x = 0$ and p is a polynomial. Example functions for f are $f(x) = \ln x$ and $f(x) = x H_0^{(1)}(x)$. It is well known that Gauss-Legendre quadrature converges exponentially for smooth integrands without singularities [134]. However, the presence of a singularity at $x = 0$ will cause the Gauss-Legendre rule to exhibit slower algebraic convergence. A simple way to overcome this issue is to use a composite Gauss-Legendre rule on a mesh geometrically graded towards the singularity. The points of the geometrically graded mesh are defined as follows.

$$x_0 := 0, \quad x_i := \sigma^{n-i} L, \quad i = 1, 2, \dots, n,$$

where n is the number of layers. On each layer $[x_{i-1}, x_i]$, $i = 1, 2, \dots, n$ we use appropriately scaled Gauss-Legendre nodes and weights. We shall adopt an hp-approach in which the number of Gauss-Legendre nodes on each layer varies across the mesh. Let the number of nodes on each layer be determined by the vector $\mathbf{p} \in \mathbb{N}_0$ which we define (as in §3.2.1) by

$$(\mathbf{p})_i := \begin{cases} p - \left\lfloor \frac{(n+1-i)}{n} p \right\rfloor, & 1 \leq i \leq n - 1, \\ p, & i = n. \end{cases}$$

We choose the integer $p = n$, i.e., equal to the number of layers in the mesh. Also, σ=0.15 was chosen as experiments revealed it to be a robust choice that achieves good convergence rates.

Figure 3.2 shows comparisons of this composite approach with standard Gauss-Legendre for $f(x) = \log(x)$ and $f(x) = x H_0^{(1)}(x)$ for $x \in [0, 1]$. It can indeed be seen that the hp-composite method has a convergence rate than far exceeds that of standard Gauss-Legendre. Close to machine precision is reached with $n = 16$ and hence $n_G = 231$ for $f(x) = \log(x)$ and with $n = 17$ for $f(x) = x H_0^{(1)}(x)$. However, it is possible to do better.

The Gauss-Legendre quadrature rule relies on the orthogonality of the Legendre polynomials, which are orthogonal to each other with weight 1. This rule, with a number of

53
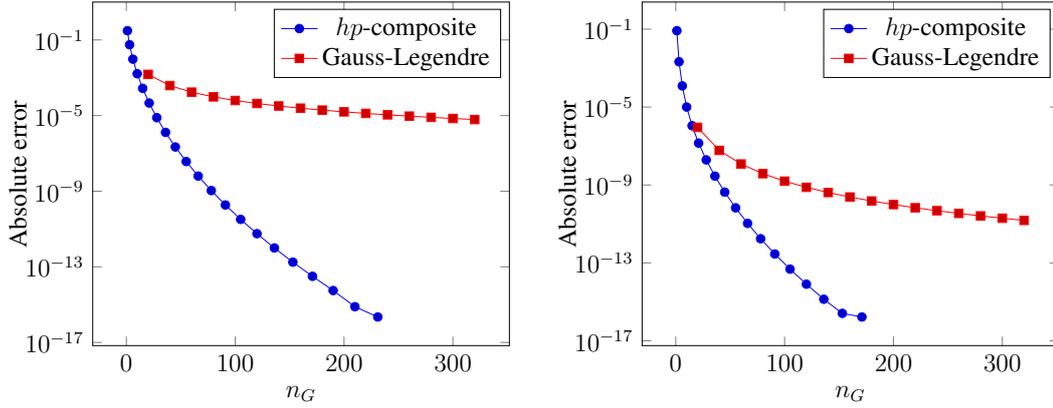
Figure 3.2: Comparison of the absolute errors in the numerical evaluation of the integrals $\int_0^1 \ln x \mathrm{d}x$ (left) and $\int_0^1 x H_0^{(1)}(x)$ (right) using the hp-composite scheme and standard Gauss-Legendre.

nodes n, integrates a polynomial of degree $2n + 1$ exactly. Since the integrand of interest here is poorly approximated by polynomials, we see poor convergence for this rule. However, if we introduce a rule which relies on polynomials orthogonal to each other with respect to the weighting function $\log(x)$, we can hope to achieve the same degree result for our integrand. Let us explain this more thoroughly.

Consider the integral

$$I[f] := \int_0^1 f(x)\mathrm{d}x, \tag{3.40}$$

for the function

$$f(x) = p(x)\log(x),$$

where $p \in \mathbb{P}$ is a polynomial. Let us define the set of functions T_n (as in [68]) for $n = 0, 1, \ldots$, by

$$T_n := \begin{cases} \{1, \log(x), x, x\log(x), \ldots, x^{l-1}\log(x), x^l\}, & n = 2l \text{ is even,} \\ \{1, \log(x), x, x\log(x), \ldots, x^{l-1}\log(x), x^l, x^l\log(x)\}, & n = 2l+1 \text{ is odd.} \end{cases} \tag{3.41}$$

They form the sequence $\{1\}, \{1, \log(x)\}, \{1, \log(x), x\}, \{1, \log(x), x, x\log(x)\}$ with the corresponding function spaces V_n being defined as

$$V_n := \mathrm{span}\left\{T_n\right\}, \quad n = 0, 1, \ldots.$$

Then the Generalised Gauss rule (GG) $Q[\cdot]$ with n nodes and weights as calculated from [68] has an associated error which satisfies

$$E^n(f) = 0 \quad \text{for all } f \in V_{2n-1}.$$

54
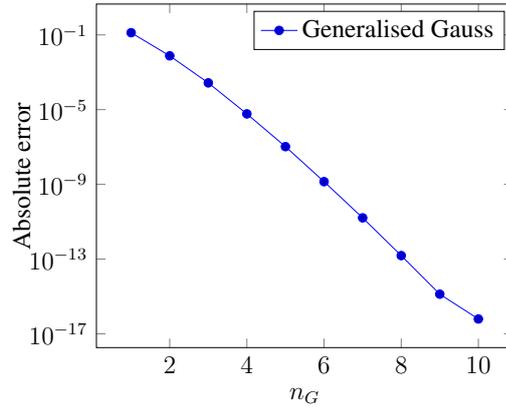
Figure 3.3: Convergence of Generalised Gauss quadrature rule for integrating $f(x) = xH_0^{(1)}(x)$ between 0 and 1.

Figure 3.3 shows the error in the approximation of $f(x) = xH_0^{(1)}(x)$ for $x \in [0,1]$. We see that machine precision is reached with 10 points, compared to 140 for the hp-composite method. The drawback to this approach is that the nodes and weights are more expensive to calculate, however to circumvent this issue, it is possible to compute the nodes and weights beforehand and store them in a lookup table.

3.2.3.2   1D integrals with near singularities

Consider the integral

$$I(\varepsilon) := \int_\varepsilon^1 p(x)f(x)\mathrm{d}x, \tag{3.42}$$

where $p(x)$ is a polynomial, $f(x)$ is a function with a logarithmic singularity at $x = 0$ and $0 < \varepsilon < 1$. If ε is small, then we expect that the presence of the nearby singularity at $x = 0$ will hinder the convergence of Gauss-Legendre quadrature. To handle such integrals, we propose a modified version of the composite rule proposed in §3.2.3.1 for the limiting case $\varepsilon = 0$.

The idea is to create a graded mesh for the interval $[0, 1]$ as in §3.2.3.1 and then simply truncate at ε to create a mesh on the interval $[\varepsilon, 1]$. On each of the elements in the new mesh, we use n_G quadrature points. This approach yields a quadrature rule which converges exponentially in p as was seen for the $\varepsilon = 0$ case. Note that this rule degenerates to the standard Gauss-Legendre rule when $\varepsilon \geq \sigma$.

To generalise the GG rule to the near singularity case, it is more complicated and expensive to calculate the nodes and weights. We shall not demonstrate how to do this here (see [68] for more details) and shall use the hp composite quadrature method for integrals with near singularities.

3.2.3.3 Singular two-dimensional integrals

For the case $[a,b] = [0,1]$, the integral (3.37) has the form

$$I[p,q,f] = \int_0^1 \int_0^1 p(x)q(y)f(x-y)\mathrm{d}y\mathrm{d}x, \tag{3.43}$$

with f being singular along the diagonal $x = y$ of the integration domain $[0,1] \times [0,1]$. A common approach here is to divide the square integration domain along this diagonal, creating two triangular domains T_1 and T_2 defined as

$$T_1 := \{(x,y) : 0 \le y \le x \le 1\}, \tag{3.44}$$
$$T_2 := \{(x,y) : 0 \le x \le y \le 1\}, \tag{3.45}$$

and to rewrite the integral (3.43) as the sum $I = I_1 + I_2$, where

$$I_1 = \int_0^1 \int_0^x p(x)q(y)f(x-y)\mathrm{d}y\mathrm{d}x, \tag{3.46}$$

$$I_2 = \int_0^1 \int_x^1 p(x)q(y)f(x-y)\mathrm{d}y\mathrm{d}x. \tag{3.47}$$

Consider first the integral I_1. We employ the Duffy transformation [41], defined here by the substitution

$$x = s, \quad y = s(1-t), \tag{3.48}$$

so that I_1 becomes

$$I_1[p,q,f] = \int_0^1 \int_0^1 p(s)q(s(1-t))f(st)s\mathrm{d}t\mathrm{d}s. \tag{3.49}$$

We see that the integration domain of (3.46) has been mapped from a triangle to a square, and that the singularity in the integrand of the transformed integral (3.49) is along the sides $s = 0$ and $s = t$ of its integration domain.

The new integral may be evaluated using a tensor product of one-dimensional GG quadrature rules so that we obtain the approximation

$$I_1[p,q,f] \approx Q_1^n[p,q,f] = \sum_{i=1}^{n_s} \sum_{j=1}^{n_t} w_i^s w_j^t p(s_i)q(s_i(1-t_j))f(s_it_j)s_i. \tag{3.50}$$

Now consider the second integral, I_2. We make the substitution

$$x = 1-s, \quad y = st + 1 - s. \tag{3.51}$$

Then I_2 becomes

$$I_2 = \int_0^1 \int_0^1 p(s)q(s(1-t))f(st)s\mathrm{d}t\mathrm{d}s \tag{3.52}$$

as for I_1, and so we may employ the same tensor product quadrature rule.

For the case when $a = 1$, the integral (3.37) takes the form

$$I[p, q, f] = \int_0^1 \int_1^b p(x)q(y)f(x - y)\mathrm{d}y\mathrm{d}x, \qquad (3.53)$$

with f being singular at the corner point $x = 1, y = 1$ of the integration domain $[0, 1] \times [1, b]$.

For such an integral, we may immediately use the product of two one-dimensional composite Gauss rules with the grading directed towards the singular point which will be located at one of the corners of the integration domain. We cannot, however, directly use the product of two one-dimensional generalised Gauss rules since for that we would require that the integrand be simultaneously singular along the entire sides $x = 1$ and $y = 1$, rather than just the one corner point. It is possible to use Duffy transformations to "smear out" the singularity along an entire side however this may lead to complications involving the introduction of near singularities which we shall not cover here. Therefore, for the sake of versatility, a tensor product of two one-dimensional composite rules is employed here for this case.

Finally, for the case when a singularity is close to the integration domain, i.e., we have an integral of the form

$$I[p, q, f] = \int_0^1 \int_{1+\varepsilon}^b p(x)q(y)f(x - y)\mathrm{d}y\mathrm{d}x, \qquad (3.54)$$

where $\varepsilon < 0.15$, a product of two composite rules is also employed. The individual one-dimensional rules are each modified according to the location of the singularity in a way identical to that discussed in §3.2.3.2.

3.2.4   Convergence and accuracy of the 2D Galerkin BEM

In §3.2.1 we presented the approximation space for a conventional hp-Galerkin BEM for the 2D transmission problem. Here we present numerical results demonstrating that the method is exponentially convergent in the polynomial degree p but we also demonstrate that, in order to maintain a prescribed error tolerance for the solution as the wavenumber k_1 increases, the number of degrees of freedom in the approximation space must increase in proportion to k_1. Hence the size of the mass matrix increases in proportion to k_1^2. This, of course, is to be expected from a conventional method. We draw attention to this fact to highlight the need for methods that do not suffer from this limitation.
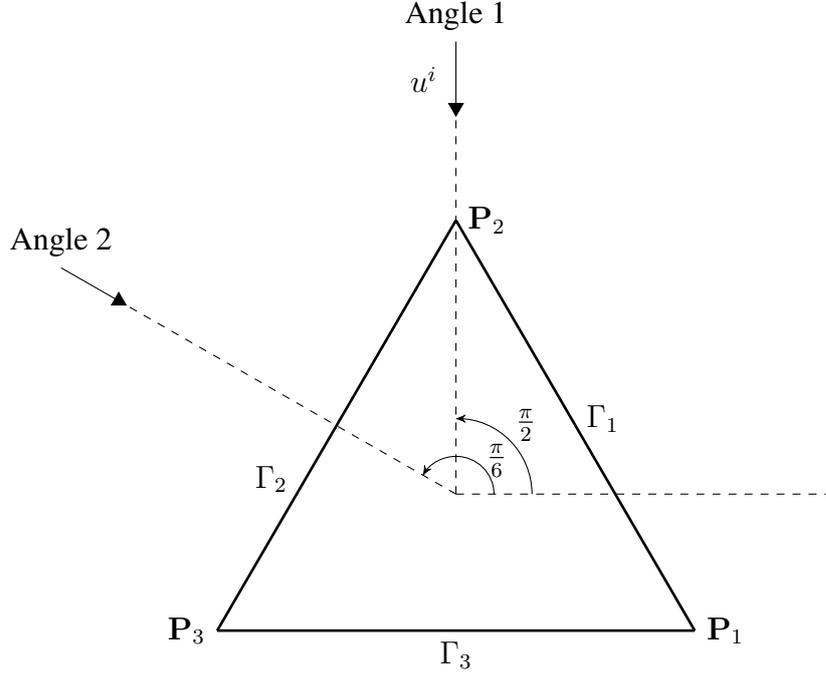
Figure 3.4: Scattering setups for triangle.

The numerical results we present shall be for the problem of the scattering by an equilateral triangle of a plane wave travelling in the direction identified with "Angle 1" in Figure 3.4. The triangle of side length 2π so that k_1 wavelengths fits along each side, and its refractive index is $n = 1.5 + 0i$. Here we also take $\alpha = 1$. In Figure 3.5 we plot the real part of the field $\partial u/\partial \mathbf{n}$ on the boundary (normalised by k_1) and in Figure 3.6 we plot the absolute value of the far-field pattern F as defined in (2.128) for the problem with $k_1 = 10$.

Figure 3.7 shows the relative L^2 error around the boundary of $\partial u/\partial \mathbf{n}$ (on a logarithmic scale) against the polynomial degree p for a range of k_1. In each case we take the "exact" reference solution to be that computed with $p = 11$. The L^2 norm is computed using the midpoint quadrature rule with 60 points per wavelength $\mathrm{Re}(\lambda_2)$, where $\lambda_2 = 2\pi/k_2$; experiments suggest that this is sufficient to achieve five digits of precision.

The linear plots in Figure 3.7 demonstrate that the error in $\partial u/\partial \mathbf{n}$ on Γ is approximately equal to $5e^{-2p}$. That is, it is exponentially decaying in p. We also observe that changing the value of k_1 has little effect on the error of the resulting approximation. This is to be expected since the mesh has been constructed so that it gets refined at a rate proportional to k_1.

Figure 3.8 compares the relative L^2 errors of $\partial u/\partial \mathbf{n}$ and the far-field pattern F for $k_1 = 5$. The L^2 norm in the far-field is approximated by the discrete L^2 norm, sampling
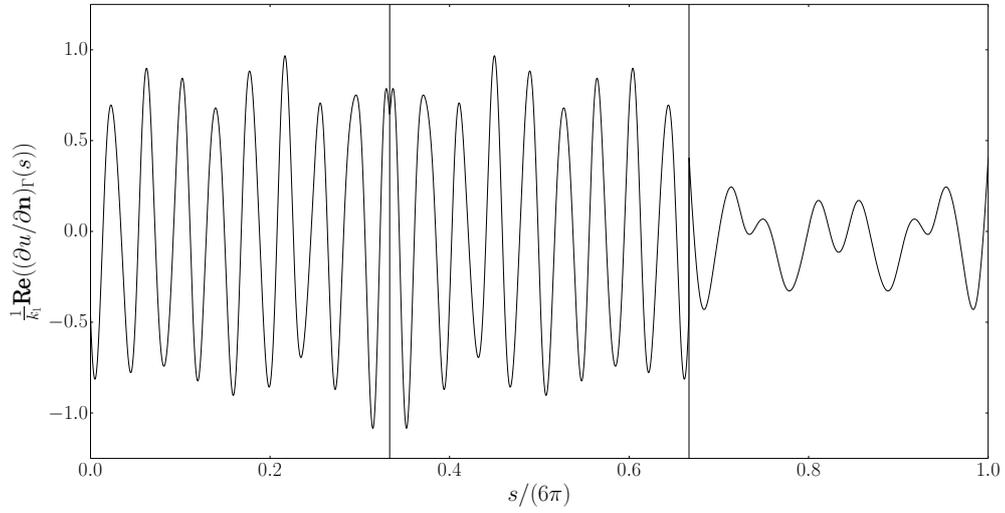
Figure 3.5: Scattering of a plane wave with $k_1 = 10$ travelling in direction defined by Angle 1 in Figure 3.4 by a triangle with $n = 1.5 + 0.05\mathrm{i}$. The boundary solution $\partial u/\partial \mathbf{n}$ (normalised by k_1).
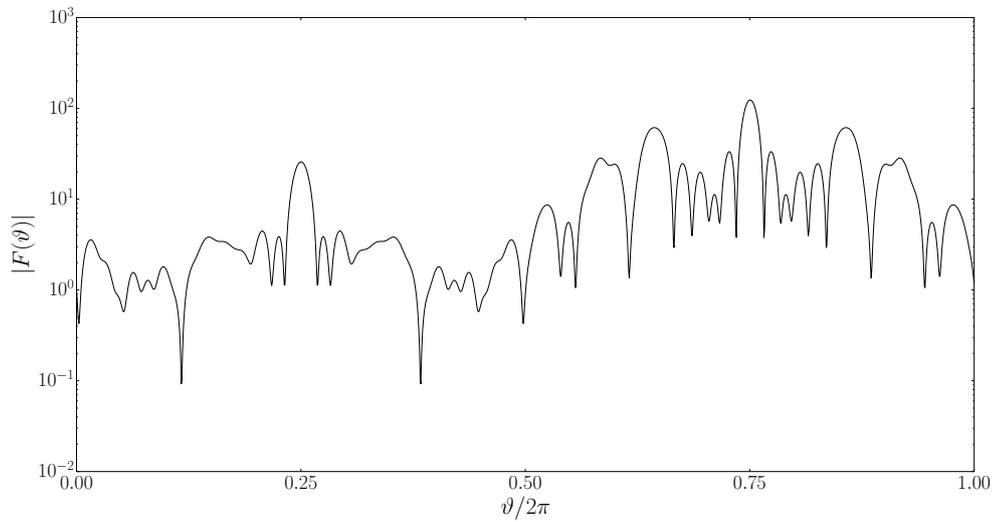
Figure 3.6: Scattering of a plane wave with $k_1 = 10$ travelling in direction defined by Angle 1 in Figure 3.4 by a triangle with $n = 1.5 + 0.05\mathrm{i}$. The absolute value of the far-field pattern F.

at $100\mathrm{Re}(\lambda_2)$ evenly spaced points on the unit circle $\mathbb{S}^1$. We see that relative errors in the far-field are often an order of magnitude, or more, smaller than the relative errors in the boundary data. Further, the trend appears to be that, the higher the value of p is, the greater the improvement in accuracy observed when mapping the boundary solution to the

59
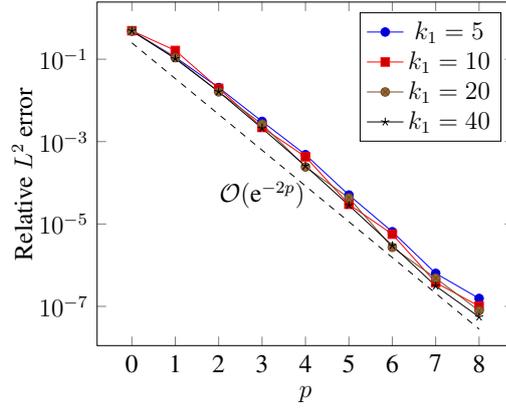
Figure 3.7: Relative L^2 errors of $\partial u/\partial \mathbf{n}$ on Γ for various values of k_1.
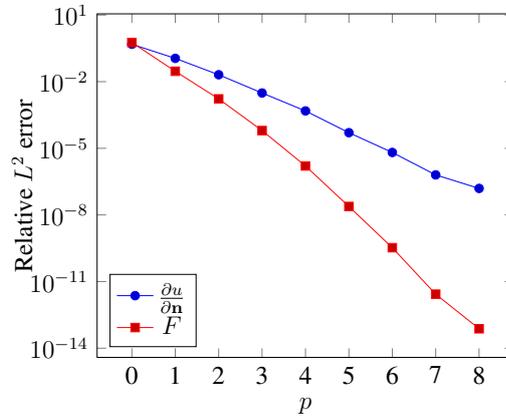
Figure 3.8: Relative L^2 errors of $\partial u \partial \mathbf{n}$ on Γ and F on the unit circle for $k_1 = 5$.

far-field.

We observe from Figure 3.8 that to achieve better than 1% error in the far-field, we require $p = 2$ in our approximation space. Employing (3.27) we see that this equates to 4.2 degrees of freedom per wavelength λ_2 (or approximately 6 per λ_1). Although this is a relatively modest memory requirement to achieve a satisfactory accuracy, it is still proportional to k_1 so that, for high enough k_1, the solution will become intractable via this method. This memory requirement is much greater for the 3D transmission problem as we shall see in the next section.

3.3 A Galerkin h-BEM for the 3D transmission problem

In this section we employ the open-source Galerkin boundary element library BEM++ [124] to approximate the solution to the 3D transmission problem (2.24)–(2.27). The purpose of this section is to demonstrate the capabilities of a state-of-the-art boundary element method

for problems of relevance to the atmospheric science community. The results presented here also serve as motivation for the development of more efficient "exact" (i.e., numerical as opposed to asymptotic) methods, which can be applied at larger size parameters.

BEM++ is a Galerkin boundary element library which has recently been developed and made open-source at http://www.bempp.org by the group of Betcke at al. [124]. The library has the capability of solving the Laplace equation, Helmholtz equation and Maxwell equations on domains with piecewise smooth Lipschitz boundaries. Here, we are interested in the Maxwell case, and, more precisely, we wish to solve the system of boundary integral equations (2.39) which is equivalent to the electromagnetic transmission problem (2.24)–(2.27). We note that the results presented here are given in more detailed form, and for a variety of scatterers (here we just focus on the hexagonal column), in the paper [53].

For solving Maxwell's equations BEM++ currently provides a single discrete approximation space, namely the space of lowest order Raviart-Thomas functions [119]. This leaves mesh refinement as the only way to increase the accuracy of the approximation. That is, we are working with a low-order h-version BEM in contrast to the high-order hp-BEM discussed in the previous section. In §3.3.1, we illustrate how the accuracy of BEM++, in approximating scattering by a hexagonal column, varies with mesh refinement, i.e., as we increase the number of mesh elements per wavelength. It is found here that (in accordance with the commonly stated rule [98]) ten boundary elements per wavelength is sufficient to achieve approximately 1% accuracy. Higher resolution meshes generate more accurate approximations at the cost of increased CPU time and memory requirements.

There are various tolerances to be set in BEM++ for the different algorithms it employs. These tolerance values were chosen after experimentation and led by examples provided on the BEM++ webpage, along with [13]. These values were chosen with the aim of ensuring that the accuracy of the solution is essentially dependent on the mesh size alone. We state the tolerances used to generate the results presented here. The adaptive cross approximation (ACA) tolerance is set as 1×10^{-5}, and the generalised minimal residual (GMRES) tolerance as 1×10^{-8}. The accuracy of the LU decomposition is chosen as 1×10^{-2}. Finally, the accuracy of the single and double regular integral quadrature was increased from the default settings by a factor of two, see [124] for details.

It should be noted that there are other ACA settings such as the maximum rank and maximum block size that may be altered to improve the performance of BEM++. It was found here, however, that for the relatively small scale computations performed here, altering these settings from their defaults made a negligible difference. For large scale computations, it is expected that adjusting these ACA settings will affect the performance of BEM++, with the optimal settings being dependent on the specifications of the computer

(or computers) being used. Finally, we note that all calculations were performed with BEM++ 2.0 which interfaces to the AHMED library [11] for $\mathcal{H}$-matrix calculations. More recent BEM++ releases come with a built-in $\mathcal{H}$-matrix implementation which may result in different memory consumption performance to the results presented here.

3.3.1 Convergence and accuracy of the 3D BEM

To get a measure of the speed, accuracy and convergence rate of BEM++, we examine the case of scattering by a hexagonal ice column with $ka = 2.5$, where a is the radius of the smallest circle which enscribes the hexagonal face, and refractive index $n = 1.311 + 2.289 \times 10^{-9}\mathrm{i}$. The aspect ratio of the hexagonal column is the ratio L/a where L is the height of the column. Throughout we take this ratio to be 2. The incident wave direction and orientation of the hexagonal ice column are both as illustrated in Figure 2.1.

We perform an experiment in which we consider uniform meshes of increasing refinement, starting from an element size of $h = \lambda_1/2.5$ and doubling the refinement for each successive numerical experiment until we reach an element size of $h = \lambda_1/40$. Such a triangulated mesh is shown for the hexagonal column in Figure 3.10. For each level of refinement, we record in Table 3.9 the memory requirement for the operator $\mathcal{S}_-$, the total run time, along with the errors in the normalised phase function P_{11} defined in (2.135). The reference solution used was calculated by the T-matrix method code of [59] which was shown there to produce accuracies of better than 0.01%. It can be seen from the table that the memory cost of BEM++ increases by approximately a factor of 5 each time the refinement of the mesh is doubled. When looking at the calculation times, it is difficult to observe a clear pattern as the mesh resolution in increased, perhaps since the calculation times are too short. But we notice that going from the $\lambda/20$ resolution to the $\lambda/40$ resolution leads to approximately a five-fold increase in computing time which is in accordance with the increase in memory requirement. We observe that the BEM is slightly better than 1% accurate at a mesh size of $h = \lambda_1/10$ when approximating P_{11}. Hence we shall use this mesh size for the remaining results presented in this section. It should be noted that typically it is advisable to grade the surface mesh towards the edges in order to optimise the number of degrees of freedom used. For scattering by perfectly conducting obstacles, this is strongly advised. However, for the dielectric ice particles of interest here, the singular behaviour of the solution near edges and vertices is weaker and so we expect that grading will give only a small advantage. Therefore we shall employ uniform meshes here for simplicity.

Next we compare the performance of a well-established T-matrix method code and BEM++. The T-matrix method was described briefly in §1.2.1 and more details can be found in the references mentioned there. The particular T-matrix code used here is that of

h	Mem. (MB)	Time (s)	P_{11} err.(%)
$\lambda/2.5$	0.0138	0.183	18.9
$\lambda/5$	0.108	0.441	9.83
$\lambda/10$	6.85	6.81	0.690
$\lambda/20$	91.9	63.0	0.212
$\lambda/40$	487	340	0.164

Table 3.9: Scattering by a hexagonal ice column with $ka = 2.5$ and $n = 1.311 + 2.289 \times 10^{-9}i$. Memory use for operator $\mathcal{R}_-$, run time, and relative error (%) in the approximation of P_{11} at each mesh refinement.
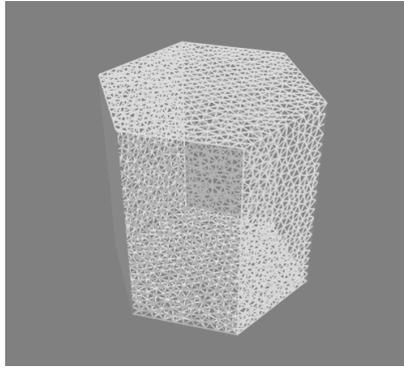
Figure 3.10: A surface mesh for the hexagonal column generated with Gmsh.

[59]. Comparisons are made for hexagonal ice columns with refractive index $n = 1.311 + 2.289 \times 10^{-9}i$ of three size parameters ($ka = 5, 10, 15$). Figure 3.11 displays P_{11} for $ka = 10$ as calculated by both the T-matrix method and BEM++. It is clear from the figure that the two methods are in excellent agreement. Table 3.12 displays some of the performance details for the two methods. BEM++, which has been optimised for computations in parallel, was run on a 4-core machine with a total of 16.4 Gigabytes of RAM, whereas the T-matrix code was run on a single core of the same machine. The memory

ka	CPU (s) (T)	CPU (s) (BEM)	Mem. (MB) (T)	Mem. (MB) (BEM)
5	319	65.5	3.81	496
10	7470	471	34.3	2420
15	27600	1720	92.3	7050

Table 3.12: CPU time and memory load utilised (given to 3 significant figures) by the T-matrix method and BEM++ to calculate the scattering properties of hexagonal ice columns of different size parameters.

requirements for BEM++ shown in the table are those required to store the four operators
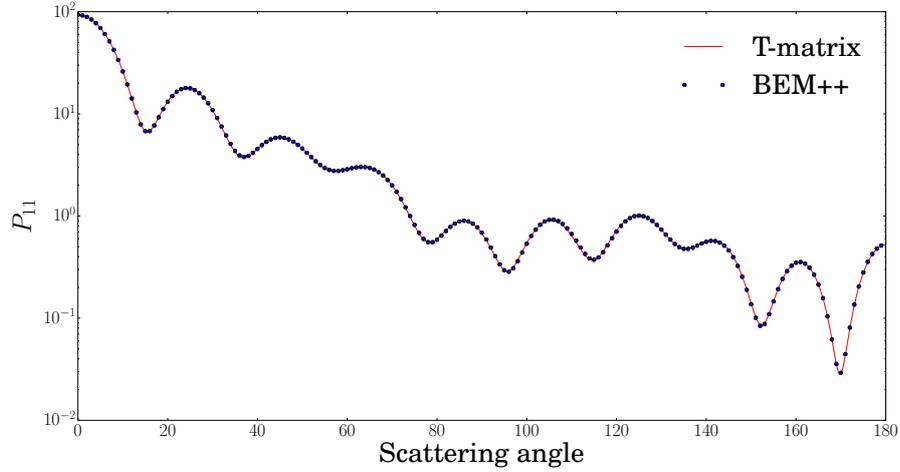
Figure 3.11: Phase function as calculated by the T-matrix method and BEM++ for $ka = 10$ and $n = 1.311 + 2.289 \times 10^{-9}$i.

$\mathcal{R}_+, \mathcal{R}_-, \mathcal{C}_+, \mathcal{C}_-$ that compose the system matrix in (2.39). Similarly, the memory requirement figures shown for the T-matrix method relate to the memory required to store the system matrix arising in that method. It is evident from the table that the memory utilised by the BEM is currently much greater than that utilised by the T-matrix code. This memory consumption is the main drawback of conventional BEMs and is the reason why BEM++ is limited to relatively small size parameters. However, we notice that due to its parallelisation and high-level implementation, BEM++ is extremely fast, with a CPU time more than 16 times faster than the T-matrix mode for $ka = 15$. Despite its lower memory requirements, it is evident that the T-matrix still becomes more memory intensive as ka grows and hence is also limited to problems with small size parameters. We found that it was also limited to $\chi = 15$ on the aforementioned computer.

Chapter 4

Asymptotic methods

High-frequency asymptotic methods are applicable when the wavelength is much smaller than a typical lengthscale a of the scattering obstacle, that is, when $ka \gg 1$. A wealth of asymptotic methods exist for wave scattering problems, many of which were mentioned in §1. Here however, we consider only two of these, both of which rely on *ray theory* which we shall review shortly in §4.1. These two methods are Geometrical Optics (GO) and the Geometrical Theory of Diffraction (GTD). Together these form the foundation for the HNA approach.

GO and GTD arise from the classical high-frequency decomposition of the total field as

$$u = u_{go} + u_d,$$

where u_{go} is the GO approximation to u and is composed of the incident, reflected and refracted fields. The remainder u_d is interpreted as the diffracted field. The motivation for this decomposition comes from asymptotic expansions for those "canonical" scattering problems for which an exact solution is known. We shall examine in more detail some such problems later, in particular, scattering by an infinite half-line and scattering by an infinite wedge.

As $ka \to \infty$, u_d vanishes and we are left with only u_{go} (except at degenerate points such as shadow boundaries - see below), hence for very high-frequency problems, the GO approximation can be sufficiently accurate for many applications. We are interested, however, in developing a method suitable for all frequencies so are required to include u_d if we wish to obtain accurate approximations.

This chapter paves the way for the construction of our HNA approximation space for scattering by a penetrable convex polygon, which we describe in §5. Here we detail the calculation of u_{go} via a beam tracing algorithm, and investigate the high-frequency phase

structure of u_d. These two tasks rely upon the analysis of certain canonical problems relevant to scattering by penetrable polygons. These are:

(i) the reflection/refraction of a plane wave at an interface between two absorbing media,

(ii) the scattering of a plane wave by an infinite penetrable wedge.

The first canonical problem is the subject of §4.3 and constitutes the majority of the discussion of the GO approximation which is the subject of §4.2 and §4.3. §4.6 discusses the natural extension of the GO approximation, namely the "Kirchhoff" or "Physical Optics" approximation, which is obtained by inputting the GO approximation on the boundary into the boundary integral representation (or typically its far-field form).

In §4.7–§4.8.2 we analyse the character of the diffracted field arising from the corners of the polygon. We begin in §4.7 by giving an overview of GTD. In order to illustrate this theory, we review the classical canonical problem of the scattering of a plane wave by an infinite half line and also by an infinite impenetrable wedge. In these cases the GTD approximations can be obtained by computing the large k asymptotics of the known exact solutions. §4.8 discusses the extension of GTD to the canonical problem of scattering by an infinite penetrable wedge for which there is no known exact solution. This discussion will lay the foundations for the development of our HNA ansatz in §5.

4.1 Ray theory

We give here a brief review of the ray theory from which both GO and the GTD stem.

We are concerned with the scattering of time-harmonic waves which are modelled by the Helmholtz equation:

$$(\nabla^2 + k^2)u = 0. \tag{4.1}$$

We first assume that all lengths have been non-dimensionalised with respect to a suitable lengthscale (such as a), so that k is a non-dimensional parameter. Ray theory assumes that u possesses the asymptotic expansion ($k \to \infty$)

$$u(\mathbf{x}, k) \sim e^{ik\psi(\mathbf{x})} \sum_{n=0}^{\infty} \left(\frac{1}{ik}\right)^n A_n(\mathbf{x}) \tag{4.2}$$

known as the *ray expansion*. Upon substitution of the ray expansion into (4.1) and equating like powers of k, one obtains the *eikonal equation* for ψ,

$$|\nabla \psi|^2 = 1, \tag{4.3}$$

and the *transport equations* for the functions $A_0, A_1, \ldots, A_n$,

$$2\nabla A_0 \cdot \nabla \psi + A_0 \nabla^2 \psi = 0, \tag{4.4}$$

$$2\nabla A_j \cdot \nabla \psi + A_j \nabla^2 \psi = -\nabla^2 A_{j-1}, \quad j > 0. \tag{4.5}$$

The eikonal equation (4.3) is a first-order, non-linear partial differential equation which can be solved using the method of characteristics. For completeness we briefly present this process which is known as "Charpit's method". Following the exposition of [113, §8.2], we first define the function

$$F(x, y, \psi, p, q) := \frac{1}{2} \left\{ \left(\frac{\partial \psi}{\partial x} \right)^2 + \left(\frac{\partial \psi}{\partial y} \right)^2 - 1 \right\}$$

$$= \frac{1}{2}(p^2 + q^2 - 1) = 0,$$

where $p = \frac{\partial \psi}{\partial x}$ and $q = \frac{\partial \psi}{\partial y}$. The general notation employed above for F is to highlight that Charpit's method can be applied more widely than for just solving the eikonal equation, see [113, §8.2] for more details. We use t to parameterise the characteristics and differentiate with respect to t to find

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\partial F}{\partial p}, \ \frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\partial F}{\partial q}, \ \frac{\mathrm{d}\psi}{\mathrm{d}t} = p\frac{\partial F}{\partial p} + q\frac{\partial F}{\partial q},$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial F}{\partial x} - p\frac{\partial F}{\partial \psi}, \ \frac{\mathrm{d}q}{\mathrm{d}t} = -\frac{\partial F}{\partial y} - q\frac{\partial F}{\partial \psi}.$$

These are known as *Charpit's equations*. For our specific F, they simplify to

$$\frac{\mathrm{d}x}{\mathrm{d}t} = p, \ \frac{\mathrm{d}y}{\mathrm{d}t} = q, \ \frac{\mathrm{d}\psi}{\mathrm{d}t} = p^2 + q^2 = 1,$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = 0, \ \frac{\mathrm{d}q}{\mathrm{d}t} = 0. \tag{4.6}$$

We see that $(p, q) = \nabla \psi$ and so both p and q remain constant on the characteristics, the distance along which is parameterised by t. The characteristics are in fact the individual *rays* composing the field. The curves normal to the rays are called *wavefronts*.

On each ray, since $\mathrm{d}p/\mathrm{d}t = \mathrm{d}q/\mathrm{d}t = 0$, we have that $p = p_0(s), q = q_0(s)$ where s parameterises a family of rays. Integrating (4.6), we obtain x, y and ψ on the rays to be

$$x(s, t) = p_0(s)t + x_0(s), \tag{4.7}$$

$$y(s, t) = q_0(s)t + y_0(s), \tag{4.8}$$

$$\psi(s, t) = t + \psi_0(s), \tag{4.9}$$

where the subscript zeros denote values on the initial curve. Equations (4.7)–(4.9) show that the rays are straight-line trajectories in a homogeneous medium. As we shall see in the next section, when an inhomogeneity is encountered the rays undergo reflection and refraction.

Now we consider the solution of the transport equations (4.4)–(4.5) along each ray. It can be shown that the Jacobian of the transition from Cartesian to ray coordinates

$$J := \left| \frac{\partial(x,y)}{\partial(s,t)} \right|$$

satisfies the equation

$$\frac{\partial J}{\partial t} = \frac{\mathrm{d}p_0}{\mathrm{d}s}\frac{\partial y}{\partial t} - \frac{\mathrm{d}q_0}{\mathrm{d}s}\frac{\partial x}{\partial t} = J\nabla^2\psi. \tag{4.10}$$

The transport equations (4.4-4.5) then become a system of ODEs along each ray:

$$2\frac{\mathrm{d}A_0}{\mathrm{d}t} + \frac{1}{J}\frac{\mathrm{d}J}{\mathrm{d}t}A_0 = 0, \tag{4.11}$$

$$2\frac{\mathrm{d}A_j}{\mathrm{d}t} + \frac{1}{J}\frac{\mathrm{d}J}{\mathrm{d}t}A_j = -\nabla^2 A_{j-1}, \ j \geq 1. \tag{4.12}$$

Integrating (4.11) and (4.12) gives

$$A_0(s,t) = A_0(s,0)\sqrt{\frac{J(s,0)}{J(s,t)}}, \tag{4.13}$$

$$A_j(s,t) = A_j(s,0)\sqrt{\frac{J(s,0)}{J(s,t)}} - \frac{1}{2}\int_0^t \sqrt{\frac{J(s,\tau)}{J(s,t)}}\nabla^2 A_{j-1}(s,\tau)\mathrm{d}\tau, \ j \geq 1. \tag{4.14}$$

When a ray of the incident plane wave strikes a planar side of a polygonal scatterer, we have that $J(s,0) = J(s,t) = 1$ (since the sides possess no curvature). Therefore, only A_0 is non-zero and $A_0(s,t) = A_0(s,0)$ for all t, that is, the ray expansion consists of only this first term which we call the GO approximation. The rays have straight line trajectories which may be tracked using the geometrical laws of reflection and refraction. The amplitude A can also be calculated simply, by using Fresnel's equations which dictate how it is affected upon reflection and refraction. These laws are obtained by considering the canonical problems of the reflection and refraction of a plane wave by an infinite interface. This chapter shall consider this canonical problem in order to derive these classical laws. We shall also consider the case when one or both of the media are absorbing. This case is much less well studied but must be analysed for the numerical algorithm proposed later in the thesis. The absorptive case leads to the generation of inhomogeneous plane waves which exhibit more complicated and sometimes counter-intuitive behaviour.

When a ray from the incident wave strikes a corner of the polygon, this gives rise to diffracted rays shed in all directions. The original ray theory described above does not take

into account these diffracted rays. However, the ray theory was systematically extended to incorporate diffracted fields by Joseph Keller in the 1960s in his famous Geometrical Theory of Diffraction. This theory proposes that the diffracted field is also composed of rays but of a different asymptotic order to those terms in the expansion (4.2). This shall be reviewed in §4.7.

4.2 Geometrical optics approximation: a beam tracing algorithm

In this section we describe the beam tracing algorithm we shall use to calculate the GO field $\mathbf{v}_{go} = (u_{go}, \partial u_{go}/\partial \mathbf{n})$ for our problem of the scattering of a plane wave, incident from the exterior domain Ω_1, by a penetrable (and potentially absorbing) obstacle denoted Ω_2.

In the GO approximation, a ray from the incident field striking a point on a smooth portion of the boundary $\Gamma = \partial \Omega_2$ gives rise to a *reflected* ray propagating back into the exterior domain Ω_1 and a *refracted* ray, propagating into the interior of the polygon Ω_2. Since we assume that Ω_2 is convex, the reflected ray propagates away to infinity without re-intersecting the boundary Γ. The refracted (or *transmitted*) ray, on the other hand, does re-intersect Γ, and if this intersection occurs on a smooth portion of Γ then further reflection/refraction occurs, with a refracted ray propagating out of the polygon into the exterior domain and an internally-reflected ray propagating back into the polygon. This internally-reflected ray can, in turn, be re-reflected/refracted, and this process continues indefinitely, giving an infinite number of internally-reflected rays, potentially all propagating in different directions.

The directions of the reflected/refracted rays and the amplitudes/phases of the fields propagating along them are governed by the well-known laws of reflection and refraction for a plane wave incident on an infinite transmission interface (i.e. the Fresnel formulae and Snell's Law). However, although these laws are completely classical in the case when both propagation media are non-absorbing (see, e.g., [19]), the generalisation to the case where one or more of the media are absorbing seems to have generated a certain amount of confusion in the literature. Also, there appears to be a phenomenon present which has been wholly unreported in the literature for the absorbing case. We provide a full derivation of the reflection/refraction laws in the general case of transmission between two absorbing media in the following section (§4.3) as well as a description of the aforementioned phenomenon.

A number of numerical algorithms have been presented for computing the GO approximation for the transmission problem using the Fresnel formulae and Snell's law (see,

e.g., [145], where the 2D problem of this paper is considered, and also [16] and [95], where a 3D analogue is studied). Many such algorithms (in particular, [145] and [95]) adopt a ray-based approach in which the incident wave is discretised into a large number of rays, each of which are traced individually as they reflect/refract within the scatterer, with the algorithm stopping after a certain (user-specified) number of internal reflections. This approach is general, in that it can be applied to smooth scatterers as well as to polygons/polyhedra. For polygons/polyhedra, however, the fact that the boundary Γ is composed of straight sides/faces means that the GO approximation consists of a collection of *beams* of rays propagating in the same direction and with the same amplitude. Each beam can be thought of as a plane wave with an associated propagation direction and amplitude, restricted to a certain subset of $\mathbb{R}^2$ (or $\mathbb{R}^3$ in the 3D case). As a result, one does not need to discretise the incident wave into a large number of rays; rather, one need only compute the propagation direction and amplitude of the plane wave associated with each beam, and record the position of the "limiting rays" which form the edges of the beam. Once the algorithm has been run once for a given geometry and incident direction, the GO approximation $\mathbf{v}_{go}$ is then readily computed at any observation point on the boundary, and for any frequency, by simply summing over the contributions from each of the beams illuminating that observation point. This is the approach we adopt in this thesis. We note that a similar approach was proposed in [52], and for the 3D problem in [16, 21].

As an illustration of the beam-tracing procedure, consider the configuration in Figure 4.1, where a penetrable hexagon is illuminated by a plane wave $u^i(\mathbf{x}) = \mathrm{e}^{\mathrm{i}k_1 \mathbf{d}^i \cdot \mathbf{x}}$ incident from the top left. In this case the incident wave strikes three of the sides of the hexagon, generating three beams of reflected rays, which propagate away to infinity, and three beams of transmitted waves, which propagate into the scatterer, as shown in Figure 4.1(a)–(c). Each of these transmitted beams has associated with it a plane wave of the form $A\mathrm{e}^{\mathrm{i}k_1(D\mathbf{d}+\mathrm{i}E\mathbf{e})\cdot\mathbf{x}}$, where the amplitude A, the propagation and decay direction vectors $\mathbf{d} \in \mathbb{R}^2$ and $\mathbf{e} \in \mathbb{R}^2$, and the constants $D > 0$ and $E \geq 0$ are determined by the reflection/refraction laws presented in section 4.3. Each beam is bounded by a pair of limiting rays, which pass through the endpoints of the side of the polygon which generated the beam. Our algorithm takes these limiting rays to be parallel to the propagation direction $\mathbf{d}$, as illustrated in Figure 4.1 (but see the discussion in Remark 4.1 below). The algorithm then tracks these limiting rays as they propagate across the interior of the scatterer, determines the points at which they re-intersect the boundary, and generates new transmitted and internally-reflected beams as appropriate, with associated plane wave directions and amplitudes again computed using the reflection/refraction laws. If the two re-intersection points of the limiting rays with the boundary lie on different sides of the polygon then

(a) Primary beams from Γ_1

(b) Primary beams from Γ_2

(c) Primary beams from Γ_3

(d) Secondary beams arising from transmitted beam in (a)

(e) Secondary beams arising from transmitted beam in (b)

(f) Secondary beams arising from transmitted beam in (c)
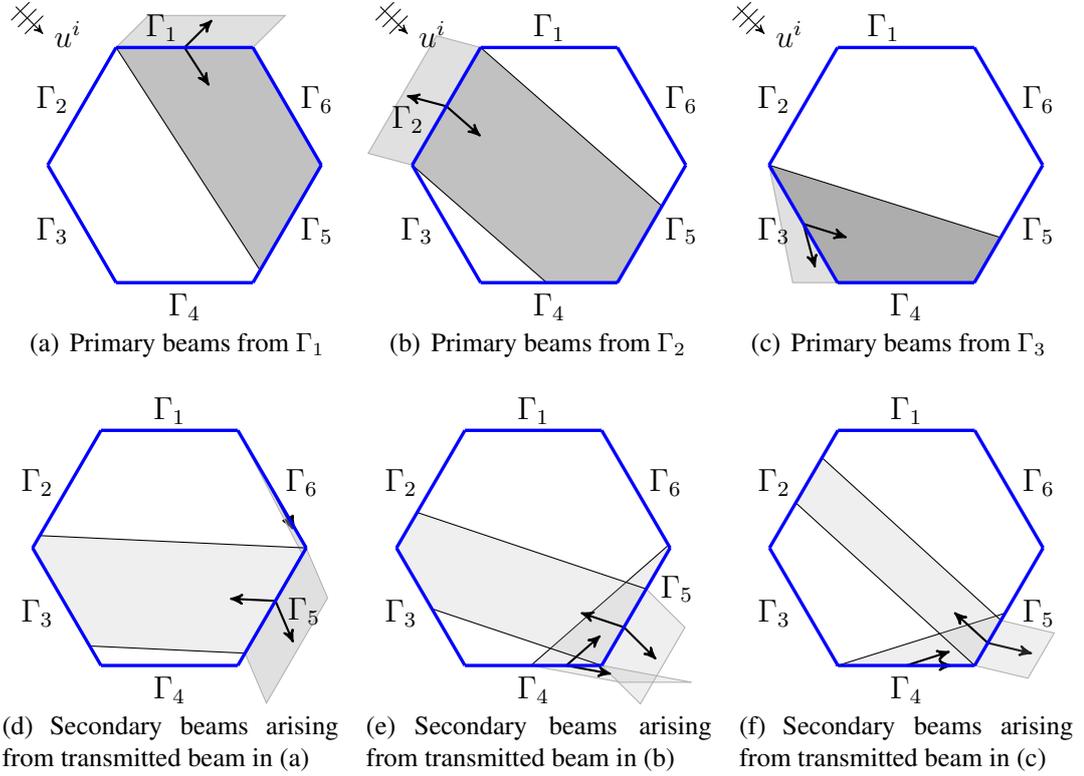
Figure 4.1: Beam tracing in a hexagon. (a)-(c) show the primary reflected and transmitted beams arising from the incidence of u^i onto sides Γ_1-Γ_3 respectively. (d)-(f) show the secondary beams arising from the internal reflection and transmission to the exterior of the primary transmitted beams in (a)-(c) respectively. Note that in each of (d) and (f) the rays associated with one of the transmitted beams point along the side - this corresponds to total internal reflection.

multiple internally-reflected beams will be produced. Figure 4.1(d)–(f) shows the two such internally-reflected beams arising from the re-reflection of each of the three beams shown in Figure 4.1(a)–(c). The algorithm continues this process of internal re-reflection until a (user-specified) stopping criterion is achieved. Some possible stopping criteria are discussed in §4.4.

Remark 4.1. *For a beam with associated plane wave $Ae^{ik_1(D\mathbf{d}+iE\mathbf{e})\cdot\mathbf{x}}$, our algorithm takes the limiting rays bounding the beam to be parallel to the propagation direction $\mathbf{d}$. This is also the choice made in [16], and it certainly seems a natural choice when the plane wave has no decay (i.e. when $E = 0$). But in the general case (in particular in an absorbing medium) it is not immediately obvious how to define the "ray direction", and hence where the "edges" of the beam should lie. What we are really asking, of course, is where the shadow boundaries between transmitted and diffracted waves lie in the related canonical diffraction problem of diffraction by an infinite absorbing transmission wedge (see sec-*

71

tion 4.8). Given the lack of an exact (or even asymptotic) solution for this infinite wedge problem (as discussed in §4.8), we cannot currently make any further comment about this. But it is interesting to note that for the related (but simpler) problem of diffraction of a general plane wave in a homogeneous absorbing medium by a sound soft knife edge, for which an exact solution is available in terms of a Fresnel integral, the correct location of the shadow boundary (defined to be the Stokes line *across which the incident field switches on/off) is* not *parallel to the real propagation vector of the plane wave beam. Rather, it is shifted somewhat in the direction of the imaginary propagation vector [12]. It would be interesting to see whether an analogous adjustment in our beam-tracing algorithm improved the accuracy of the GO approximation, but we leave further investigation of this for future work.*

4.3   Reflection and refraction at a planar interface

In this section we consider the reflection and refraction of a plane wave at an interface between two differing media. This is one of the most well-studied problems in physics and has been understood, to a greater or lesser extent, for well over a thousand years. The earliest known manuscript on the topic is Ibn Sahl's treatise *On Burning Mirrors and Lenses* of 984 AD [122], in which he alludes to work from earlier Greek texts. The laws of reflection and refraction were later rediscovered by Willebrord Snellius (1580-1626) and René Descartes (1596-1650). The laws we refer to, commonly called the *law of reflection* and *Snell's law* or sometimes the *Snell-Descartes law*, dictate the directions of the reflected and refracted plane waves, respectively. The discovery of the formulae for the relation between the amplitudes of these waves to that of the incident wave is attributed to Augustin-Jean Fresnel (1788-1827), and hence these formulae are often called the *Fresnel equations*.

The classical cases, discussed in almost every physics text book (e.g., [19, 76, 130]), are:

 (i) when both media are non-absorbing (the refractive indices of both media are real),

 (ii) when the second medium alone is absorbing (the refractive index of the second medium is complex).

We remark that in both of these cases the incident wave in the first medium is an unbounded homogeneous (see §4.3.1) plane wave. This is clearly an idealisation. Nevertheless, such a wave in a non-absorbing medium is often used as the incident field in scattering problems and can be justified as being a good approximation to the wavefronts produced by a point source located far from the interface. Also, in both these cases, there is the possibility

of inhomogeneous (see §4.3.1) plane waves being generated in the second medium. Such waves possess exponential decay, the direction of which may be chosen such that the wave decays as it progresses into the second medium.

The cases which have been avoided in the classical physics books are:

(iii) when both media are absorbing,

(iv) when the first medium alone is absorbing.

An issue with these final two cases is that we must now consider unbounded inhomogeneous plane waves in the upper medium which, as Clemmow states in [34], "represent a grosser violation of physical reality [than that of an unbounded homogeneous plane wave], since their exponential decay in one direction is matched by exponential growth in the opposite direction; it is unacceptable to allow the latter to develop without limit".

This "gross violation" of reality has still not prevented many authors from attempting to solve this problem in cases (iii) and (iv). Papers published on this subject include [32, 42, 55, 96, 116, 149]. Many of these are similar in their methodology to Pincherle's approach in [116], although few cite his work. Some authors differ in their derived formulae; and none (to my knowledge) provide experimental verification of their formulae. The differences result from the physical constraints imposed on the refracted wave. As we shall see in what is to follow, the refracted wave is, in general, an inhomogeneous plane wave with a propagation direction $\mathbf{d}^t$ and a decay direction $\mathbf{e}^t$. These directions are not defined uniquely from the formulae since there is a sign choice to be made stemming from a square root. But once a sign choice is made for one of the vectors $\mathbf{d}^t$ and $\mathbf{e}^t$, the other is specified immediately from the preceding formulae.

Pincherle justifies a particular sign choice for $\mathbf{d}^t$ (and hence for $\mathbf{e}^t$) based on energy flow and Poynting's vector, however this leads to some peculiar discontinuities in the wave behaviour as certain parameters of the problem are varied. Dupertuis et al. [42] follow from Pincherle but extend it to the three-dimensional electromagnetic problem. Here, we shall also follow Pincherle but in addition perform numerical experiments to test this sign choice. These experiments entail the computation of "exact" solutions to wave scattering by absorbing polygons and the comparison of the solution on the boundary to that obtained using a beam tracing algorithm employing Snell's Law and the Fresnel equations to be derived. We shall find that Pincherle's choice leads to poor agreement with the exact solution when the wave is close to total internal reflection (a phenomenon which cannot completely occur when absorbing media are involved). In fact, the solution is almost completely out of phase. When this occurs, we find that making the other, somewhat counterintuitive, sign choice provides excellent agreement with the exact solution.

In addition, it should be pointed out that Chang et al. in [32] employ a similar approach to Pincherle but choose $\mathbf{e}^t$ in such a way that the refracted waves fail to satisfy the governing equation. This is explained in §4.3.3 in detail. We immediately disregard such an approach since a wave that does not satisfy the governing equation cannot be the physically correct solution.

We shall begin by discussing the characterstics of inhomogeneous plane waves in an absorbing medium. Then we shall analyse the most general problem of the reflection and refraction of an inhomogeneous plane wave at the interface between two absorbing media. Finally, we shall look in more detail at the two cases of specific relevance for our problem of scattering by an absorbing particle embedded in a non-absorbing medium. These are the cases (ii) and (iv) given earlier.

4.3.1 Plane wave propagation in an absorbing medium

We consider time-harmonic scalar waves modelled by the Helmholtz equation

$$\Delta u + k^2 u = 0, \tag{4.15}$$

with a complex wavenumber k. We shall write

$$k = k_0(\mu + \mathrm{i}\xi),$$

where $k_0 > 0$ is a reference real wavenumber, and $\mu + \mathrm{i}\xi$ is the refractive index of the medium. We shall assume throughout that $\mu > 0$ and $\xi \geq 0$; the case $\xi = 0$ corresponds to a non-absorbing medium, and the case $\xi > 0$ corresponds to an absorbing medium. We consider solutions of (4.15) of the form

$$u(\mathbf{x}) = A\mathrm{e}^{\mathrm{i}k_0 \mathbf{V} \cdot \mathbf{x}}, \qquad \mathbf{x} \in \mathbb{R}^2, \tag{4.16}$$

which represents a plane wave with complex amplitude A and complex propagation vector $\mathbf{V}$. It is convenient to split $\mathbf{V}$ into real and imaginary components, writing

$$\mathbf{V} = D\mathbf{d} + \mathrm{i}E\mathbf{e},$$

where $\mathbf{d}$ and $\mathbf{e}$ are real unit vectors and D, E are real scalars, after which (4.16) becomes

$$u(\mathbf{x}) = A\exp\{\mathrm{i}k_0(D\mathbf{d} + \mathrm{i}E\mathbf{e}) \cdot \mathbf{x}\}. \tag{4.17}$$

Note that the vectors $\mathbf{d}$ and $\mathbf{e}$ are normal to the planes of constant phase and constant amplitude of u, respectively. Without loss of generality we may assume that $\mathbf{d} \cdot \mathbf{e} \geq 0$. If $\mathbf{d} = \mathbf{e}$, then the wave is called *homogeneous*, otherwise it is *inhomogeneous*. In order to

find the relationship between $\mathbf{d}, \mathbf{e}, D, E$ and the components μ, ξ of the refractive index, we substitute (4.17) into the Helmholtz equation (4.15), which, after equating the real and imaginary parts, gives

$$D^2 - E^2 = \mu^2 - \xi^2, \tag{4.18}$$

$$(D\mathbf{d}) \cdot (E\mathbf{e}) = \mu\xi. \tag{4.19}$$

We note from (4.18)–(4.19) that we cannot have $D = 0$; otherwise (4.19) would imply that $\xi = 0$ (since $\mu > 0$ by assumption), and (4.18) would then give the contradiction $-E^2 = \mu^2$. Without loss of generality we may assume that $D > 0$ (we can multiply both $\mathbf{d}$ and $\mathbf{e}$ by minus one if necessary). Now, if $\xi > 0$ then $\mu\xi > 0$ and (4.19) implies that $\mathbf{d} \cdot \mathbf{e} > 0$ and $DE > 0$, so that $E > 0$ too. On the other hand, if $\xi = 0$, then (4.19) reduces to $(D\mathbf{d}) \cdot (E\mathbf{e}) = 0$. Since the real component $D\mathbf{d}$ is non-zero, this implies that either $E\mathbf{e} = 0$ (i.e. $E = 0$ and $D = \mu$) or $\mathbf{d}$ and $\mathbf{e}$ are perpendicular. In the latter case there are an infinite family of pairs (D, E) which satisfy (4.18). Without loss of generality we can, in this case, assume that $E > 0$ (we can multiply $\mathbf{e}$ by minus one if necessary).

To summarise, we have shown that if (4.17) is a solution of (4.15) then $\mathbf{d}, \mathbf{e}, D$ and E must satisfy (4.18)–(4.19) and without loss of generality we may assume that $D > 0$, $E \geq 0$, and $\mathbf{d} \cdot \mathbf{e} \geq 0$. Under these assumptions, the wave (4.16) propagates in the direction of $\mathbf{d}$, while decaying in the direction $\mathbf{e}$; in fact, we note that D and E are interpreted by some authors as the real and imaginary parts of an "apparent refractive index" [32, 145].

4.3.2 An interface between two media with arbitrary absorption

We now consider the canonical problem of the reflection/refraction of an incident plane wave of the general form (4.17) propagating in a medium with refractive index $\mu_1 + \mathrm{i}\xi_1$ at a planar interface with a second medium with refractive index $\mu_2 + \mathrm{i}\xi_2$. We assume that in the first medium the field takes the form $u = u^i + u^r$, where u^i is the incident plane wave and u^r is a reflected plane wave, and that in the second medium the field takes the form $u = u^t$, where u^t is a transmitted plane wave. We also assume that both the total field u and its normal derivative are continuous across the interface, which implies that, on the interface,

$$u^i + u^r = u^t \quad \text{and} \quad \frac{\partial u^i}{\partial \mathbf{n}} + \frac{\partial u^r}{\partial \mathbf{n}} = \alpha \frac{\partial u^t}{\partial \mathbf{n}}, \tag{4.20}$$

where $\mathbf{n}$ is a vector normal to the interface. We write the waves u^i, u^r and u^t in the general form (4.17) as:

$$
\begin{aligned}
u^i &= A^i \exp\{ik_0(D_i\mathbf{d}^i + iE_i\mathbf{e}^i) \cdot \mathbf{x}\}, \\
u^r &= A^r \exp\{ik_0(D_i\mathbf{d}^r + iE_i\mathbf{e}^r) \cdot \mathbf{x}\}, \\
u^t &= A^t \exp\{ik_0(D_t\mathbf{d}^t + iE_t\mathbf{e}^t) \cdot \mathbf{x}\},
\end{aligned}
\tag{4.21}
$$

where we have assumed a priori the same "apparent refractive index" for the reflected wave as for the incident wave. Given the parameters A^i, $\mathbf{d}^i$, $\mathbf{e}^i$, D_i and E_i describing the incident wave, we wish to determine the parameters A^r, A^t, $\mathbf{d}^r$, $\mathbf{e}^r$, $\mathbf{d}^t$, $\mathbf{e}^t$, D_t and E_t determining the reflected and transmitted waves.

The geometry of the problem is illustrated in Figure 4.2. The real and imaginary components of the direction vectors have been drawn on separate diagrams for clarity, but it should be kept in mind that the complex incident direction vector is $\mathbf{d}^i + i\mathbf{e}^i$ and that the reflected and transmitted direction vectors are $\mathbf{d}^r + i\mathbf{e}^r$ and $\mathbf{d}^t + i\mathbf{e}^t$, respectively.

(a) Real components of direction vectors.   (b) Imaginary components of direction vectors.
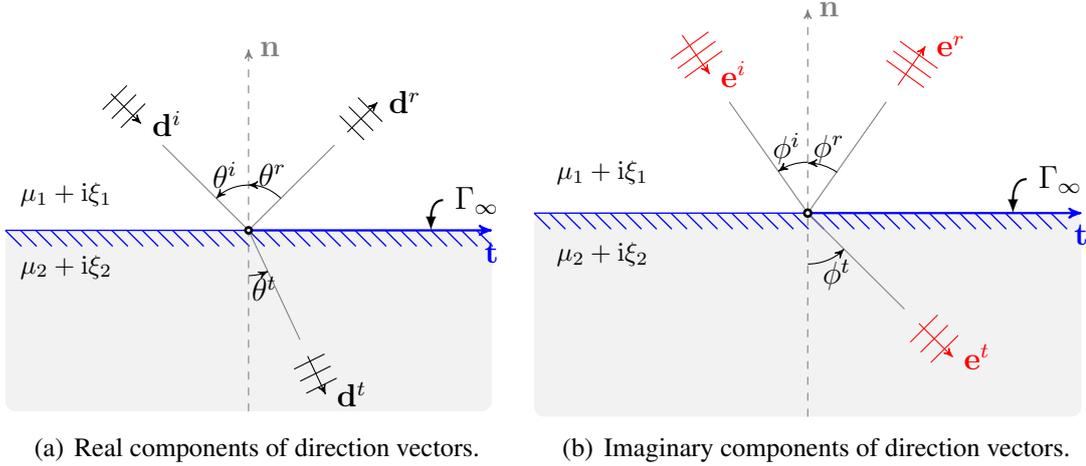
Figure 4.2: Refraction and reflection of light at the interface Γ_∞

The boundary conditions (4.20) imply that the spatial variation of the incident and reflected fields must match at the interface [73, p.217]. This leads directly to the fact that the that the real and imaginary components of the incident and reflected direction vectors satisfy the *specular reflection law* ("angle of reflection equals angle of incidence"), which can be stated in vector notation as

$$
\begin{aligned}
\mathbf{d}^r &= \mathbf{d}^i - 2(\mathbf{d}^i \cdot \mathbf{n})\mathbf{n}, \\
\mathbf{e}^r &= \mathbf{e}^i - 2(\mathbf{e}^i \cdot \mathbf{n})\mathbf{n},
\end{aligned}
\tag{4.22}
$$

or, in the notation of Figure 4.2, simply as

$$
\theta^i = \theta^r, \quad \phi^i = \phi^r.
$$

76

The tangential components of the incident and transmitted direction vectors are related via Snell's law. The Fresnel formulae relate the amplitudes of the reflected and transmitted waves to that of the incident wave. To derive these relationships we first substitute the representations (4.21) into the boundary conditions (4.20). Eliminating $\mathbf{d}^r$ and $\mathbf{e}^r$ using the law of reflection (4.22) gives, for any $\mathbf{x} \in \Gamma_\infty$,

$$\left(A^i + A^r \exp\{-2\mathrm{i}k_0(D_i\mathbf{d}^i \cdot \mathbf{n} + E_i\mathbf{e}^i \cdot \mathbf{n})\mathbf{n} \cdot \mathbf{x}\}\right)\exp\{\mathrm{i}k_0(D_i\mathbf{d}^i + \mathrm{i}E_i\mathbf{e}^i) \cdot \mathbf{x}\}$$
$$= A^t \exp\{\mathrm{i}k_0(D_t\mathbf{d}^t + \mathrm{i}E_t\mathbf{e}^t) \cdot \mathbf{x}\}, \tag{4.23}$$

and

$$v^i\left(A^i - A^r \exp\{-2\mathrm{i}k_0(D_i\mathbf{d}^i \cdot \mathbf{n} + E_i\mathbf{e}^i \cdot \mathbf{n})\mathbf{n} \cdot \mathbf{x}\}\right)\exp\{\mathrm{i}k_0(D_i\mathbf{d}^i + \mathrm{i}E_i\mathbf{e}^i) \cdot \mathbf{x}\}$$
$$= \alpha v^t A^t \exp\{\mathrm{i}k_0(D_t\mathbf{d}^t + \mathrm{i}E_t\mathbf{e}^t) \cdot \mathbf{x}\}, \tag{4.24}$$

where $v^i = D_i\mathbf{d}^i \cdot \mathbf{n} + \mathrm{i}E_i\mathbf{e}^i \cdot \mathbf{n}$ and $v^t = D_t\mathbf{d}^t \cdot \mathbf{n} + \mathrm{i}E_t\mathbf{e}^t \cdot \mathbf{n}$.

To derive Snell's law we write $\mathbf{x} \in \Gamma_\infty$ as

$$\mathbf{x} = \mathbf{X} + s\mathbf{t}, \tag{4.25}$$

where $\mathbf{X}$ is an arbitrary reference point on Γ_∞, $s \in \mathbb{R}$ and $\mathbf{t}$ is the unit tangent vector to Γ_∞ defined as $\mathbf{t} = (n_2, -n_1)$, where $\mathbf{n} = (n_1, n_2)$. Substituting (4.25) into (4.23) and rearranging gives

$$A^t = \exp\{\mathrm{i}k_0 s(D_i\mathbf{d}^i + \mathrm{i}E_i\mathbf{e}^i - D_t\mathbf{d}^t - \mathrm{i}E_t\mathbf{e}^t) \cdot \mathbf{t}\}\times$$
$$\exp\{\mathrm{i}k_0(D_i\mathbf{d}^i + \mathrm{i}E_i\mathbf{e}^i - D_t\mathbf{d}^t - \mathrm{i}E_t\mathbf{e}^t) \cdot \mathbf{X}\}\left(A^i + A^r\exp\{-2\mathrm{i}k_0(D_i\mathbf{d}^i \cdot \mathbf{n} + \mathrm{i}E_i\mathbf{e}^i \cdot \mathbf{n})\mathbf{n} \cdot \mathbf{X}\}\right).$$
$$\tag{4.26}$$

Since this must hold for all $\mathbf{x} \in \Gamma_\infty$, i.e. for all $s \in \mathbb{R}$, the argument of the first exponential factor on the right-hand side must be equal to zero, i.e.

$$D_i\mathbf{d}^i \cdot \mathbf{t} + \mathrm{i}E_i\mathbf{e}^i \cdot \mathbf{t} - D_t\mathbf{d}^t \cdot \mathbf{t} - \mathrm{i}E_t\mathbf{e}^t \cdot \mathbf{t} = 0. \tag{4.27}$$

Comparing real and imaginary components of (4.27) then yields the vector form of *Snell's Law*:

$$\begin{aligned} D_i\mathbf{d}^i \cdot \mathbf{t} &= D_t\mathbf{d}^t \cdot \mathbf{t}, \\ E_i\mathbf{e}^i \cdot \mathbf{t} &= E_t\mathbf{e}^t \cdot \mathbf{t}, \end{aligned} \tag{4.28}$$

which can also be written in more classical form in terms of the notation of Figure 4.2 as

$$D_i \sin\theta^i = D_t \sin\theta^t, \qquad E_i \sin\phi^i = E_t \sin\phi^t.$$

To derive the Fresnel formulae, we note that, given the incident amplitude A^i, equations (4.23) and (4.24) are simultaneous equations in A^r and A^t which can be solved to give the reflection and transmission coefficients

$$R := \frac{A^r}{A^i} = \frac{v_i - \alpha v_t}{v_i + \alpha v_t} \exp\{2\mathrm{i}k_0(D_i \mathbf{d}^i \cdot \mathbf{n} + \mathrm{i}E_i \mathbf{e}^i \cdot \mathbf{n})\mathbf{n} \cdot \mathbf{X}\}, \qquad (4.29)$$

$$T := \frac{A^t}{A^i} = \frac{2v_i}{v_i + \alpha v_t} \exp\{\mathrm{i}k_0(D_i \mathbf{d}^i + \mathrm{i}E_i \mathbf{e}^i - D_t \mathbf{d}^t - \mathrm{i}E_t \mathbf{e}^t) \cdot \mathbf{X}\}, \qquad (4.30)$$

respectively, where, as above, $\mathbf{X}$ is an arbitrary reference point on Γ_∞. If the origin of our coordinate system lies on Γ_∞ then we may take $\mathbf{X} = \mathbf{0}$, giving

$$R = \frac{v_i - \alpha v_t}{v_i + \alpha v_t}, \qquad\qquad T = \frac{2v_i}{v_i + \alpha v_t}, \qquad (4.31)$$

along with the classical relationship $1 + R = T$.

We recall from §4.3.1 that D_t and E_t must satisfy the equations

$$D_t^2 - E_t^2 = \mu_2^2 - \xi_2^2, \qquad (4.32)$$
$$D_t E_t \mathbf{d}^t \cdot \mathbf{e}^t = \mu_2 \xi_2. \qquad (4.33)$$

Equation (4.33) can be written in terms of tangential and normal components as

$$D_t E_t [(\mathbf{d}^t \cdot \mathbf{t})(\mathbf{e}^t \cdot \mathbf{t}) + (\mathbf{d}^t \cdot \mathbf{n})(\mathbf{e}^t \cdot \mathbf{n})] = \mu_2 \xi_2. \qquad (4.34)$$

After rearranging (4.34), squaring, and writing the normal components in terms of the tangential components (using the fact that $(\mathbf{d}^t \cdot \mathbf{t})^2 + (\mathbf{d}^t \cdot \mathbf{n})^2 = (\mathbf{e}^t \cdot \mathbf{t})^2 + (\mathbf{e}^t \cdot \mathbf{n})^2 = 1$), Snell's law implies that

$$(D_t^2 - \tilde{D}_i^2)(E_t^2 - \tilde{E}_i^2) = (\mu_2 \xi_2 - \tilde{D}_i \tilde{E}_i)^2,$$

where $\tilde{D}_i := D_i \mathbf{d}^i \cdot \mathbf{t}$ and $\tilde{E}_i := E_i \mathbf{e}^i \cdot \mathbf{t}$. Finally, using (4.32) to eliminate E_t, we arrive at a quadratic equation satisfied by D_t^2,

$$D_t^4 + D_t^2[\xi_2^2 - \mu_2^2 - \tilde{E}_i^2 - \tilde{D}_i^2] + \tilde{D}_i^2(\mu_2^2 - \xi_2^2) - (\mu_2 \xi_2)^2 + 2\mu_2 \xi_2 \tilde{D}_i \tilde{E}_i = 0, \qquad (4.35)$$

and the quadratic formula yields the solutions of (4.35) as

$$D_t^2 = \frac{1}{2}\left(\mu_2^2 - \xi_2^2 + \tilde{D}_i^2 + \tilde{E}_i^2 \pm \sqrt{(\mu_2^2 - \xi_2^2 - \tilde{D}_i^2 + \tilde{E}_i^2)^2 + 4(\tilde{D}_i \tilde{E}_i - \mu_2 \xi_2)^2}\right). \qquad (4.36)$$

Similar equations have been derived in [32] and [145]; however, the correct sign to choose in (4.36) is not discussed in these references. We claim that the positive square root should

be taken in (4.36) for consistency with Snell's law. To justify this statement, we note that Snell's law trivially implies the inequalities

$$D_t^2 \geq \tilde{D}_i^2, \tag{4.37}$$

$$E_t^2 \geq \tilde{E}_i^2. \tag{4.38}$$

We can rearrange (4.36) to give

$$D_t^2 - \tilde{D}_i^2 = \frac{1}{2}(a \pm \sqrt{a^2 + b^2}), \tag{4.39}$$

$$E_t^2 - \tilde{E}_i^2 = \frac{1}{2}(-a \pm \sqrt{a^2 + b^2}), \tag{4.40}$$

where $a := \mu_2^2 - \xi_2^2 - \tilde{D}_i^2 + \tilde{E}_i^2$ and $b := 2(\tilde{D}_i\tilde{E}_i - \mu_2\xi_2)$. Then if $a < 0$, it is clear from (4.39) that we must take the positive square root in order to satisfy (4.37). If $a > 0$, it is clear from (4.40) that we must take the positive square root in order to satisfy (4.38). If $a = 0$, then we must take the positive square root in order to satisfy both (4.37) and (4.38), unless of course $b = 0$ too, in which case the sign choice is immaterial.

Having justitifed the choice of the positive square root in (4.36), we can state the formulae for D_t and E_t:

$$D_t = \sqrt{\frac{1}{2}\left(\mu_2^2 - \xi_2^2 + \tilde{D}_i^2 + \tilde{E}_i^2 + \sqrt{(\mu_2^2 - \xi_2^2 - \tilde{D}_i^2 + \tilde{E}_i^2)^2 + 4(\tilde{D}_i\tilde{E}_i - \mu_2\xi_2)^2}\right)},$$

$$E_t = \sqrt{D_t^2 + \xi_2^2 - \mu_2^2},$$

$$\tag{4.41}$$

where the non-negative square root is taken in both equations. Snell's law provides a formula for the tangential components of the transmitted direction vectors $\mathbf{d}^t$ and $\mathbf{e}^t$. The fact that $\mathbf{d}^t$ and $\mathbf{e}^t$ are unit vectors allows us to write

$$\mathbf{d}^t = (\mathbf{d}^t \cdot \mathbf{t})\mathbf{t} + (\mathbf{d}^t \cdot \mathbf{n})\mathbf{n} = (\mathbf{d}^t \cdot \mathbf{t})\mathbf{t} \pm \sqrt{1 - (\mathbf{d}^t \cdot \mathbf{t})^2}\mathbf{n}, \tag{4.42}$$

$$\mathbf{e}^t = (\mathbf{e}^t \cdot \mathbf{t})\mathbf{t} + (\mathbf{e}^t \cdot \mathbf{n})\mathbf{n} = (\mathbf{e}^t \cdot \mathbf{t})\mathbf{t} \pm \sqrt{1 - (\mathbf{e}^t \cdot \mathbf{t})^2}\mathbf{n}, \tag{4.43}$$

so that the normal components are specified up to sign. The need to make a sign choice in (4.42) and (4.43) is alluded to in [42, p. 1163], but a clear prescription of which sign to take is not provided there. In fact, this sign choice is the pivotal part in a succesful ray tracing routine which has so far (to the author's knowledge) not been properly discussed and tested. Here we shall provide two plausible physical arguments to justify particular sign choices. In certain situations the two arguments are incompatible in that they produce different sign choices. We shall test each sign choice in numerical experiments in order

to heuristically establish which sign choice should be employed. In fact, we shall see that there exist two different regimes, within each of which it is appropriate to make a different sign choice.

The two physical arguments are as follows. Firstly, one can argue that the average flow of energy must be directed from the first medium into the second (as in [116] for the EM case). Secondly, one can argue that the transmitted wave must decay as $y \to -\infty$ in the second medium. (The latter argument is often used to decide $\mathbf{e}^t$ in the case of total internal reflection in physics textbooks when two non-absorbing media are considered.) Having described the two arguments in detail, we shall then perform numerical experiments in which the two different arguments are implemented in a beam tracing algorithm and the solution of the boundary compared to that obtained with a BEM. We shall find that each physical argument appears to have its regime of validity which depends on the incident angle. It is at present not clear why the transition between these two regimes should occur.

4.3.3 Sign choice

The multi-valuedness of equations (4.42)–(4.43) indicates that a sign choice must be made. We are required only to choose the sign in one of (4.42) and (4.43) since the correct sign in the other equation follows immediately from (4.34). We state the two different sign choices and the resulting equations for $\mathbf{d}^t$ and $\mathbf{e}^t$, and then we provide a physical argument to justify each choice and derive the given formulae. The two choices, which we refer to as GO1 and GO2, are as follows.

GO1. Force $\mathbf{d}^t$ to point into the second medium, i.e., force $\operatorname{sgn}(\mathbf{d}^t \cdot \mathbf{n}) = \operatorname{sgn}(\mathbf{d}^i \cdot \mathbf{n})$. This leads to the following equations for $\mathbf{d}^t$ and $\mathbf{e}^t$.

$$\mathbf{d}^t = \frac{D_i}{D_t}(\mathbf{d}^i \cdot \mathbf{t})\mathbf{t} + \operatorname{sgn}(\mathbf{d}^i \cdot \mathbf{n})\sqrt{1 - \left(\frac{D_i}{D_t}\right)^2 (\mathbf{d}^i \cdot \mathbf{t})^2}\mathbf{n}, \tag{4.44}$$

$$\mathbf{e}^t = \frac{E_i}{E_t}(\mathbf{e}^i \cdot \mathbf{t})\mathbf{t} + \frac{1}{D_t E_t \mathbf{d}^t \cdot \mathbf{n}} \left(\mu_2 \xi_2 - D_i E_i (\mathbf{d}^i \cdot \mathbf{t})(\mathbf{e}^i \cdot \mathbf{t})\right)\mathbf{n}, \quad E_t \mathbf{d}^t \cdot \mathbf{n} \neq 0. \tag{4.45}$$

When $E_t = 0$, we may choose $\mathbf{e}^t$ arbitrarily ($\mathbf{e}^t = \mathbf{d}^t$, for example). When $\mathbf{d}^t \cdot \mathbf{n} = 0$, replace (4.45) with

$$\mathbf{e}^t = \frac{E_i}{E_t}(\mathbf{e}^i \cdot \mathbf{t})\mathbf{t} + \operatorname{sgn}(\mathbf{d}^i \cdot \mathbf{n})\sqrt{1 - \left(\frac{E_i}{E_t}\right)^2 (\mathbf{e}^i \cdot \mathbf{t})^2}\mathbf{n}. \tag{4.46}$$

GO2. Force $\mathbf{e}^t$ to point into the second medium, i.e., force $\mathrm{sgn}(\mathbf{e}^t \cdot \mathbf{n}) = \mathrm{sgn}(\mathbf{d}^i \cdot \mathbf{n})$. This leads to the following equations for $\mathbf{d}^t$ and $\mathbf{e}^t$.

$$\mathbf{e}^t = \frac{E_i}{E_t}(\mathbf{e}^i \cdot \mathbf{t})\mathbf{t} + \mathrm{sgn}(\mathbf{d}^i \cdot \mathbf{n})\sqrt{1 - \left(\frac{E_i}{E_t}\right)^2 (\mathbf{e}^i \cdot \mathbf{t})^2}\,\mathbf{n}, \qquad (4.47)$$

$$\mathbf{d}^t = \frac{D_i}{D_t}(\mathbf{d}^i \cdot \mathbf{t})\mathbf{t} + \frac{1}{D_t E_t \mathbf{e}^t \cdot \mathbf{n}}\left(\mu_2 \xi_2 - D_i E_i (\mathbf{d}^i \cdot \mathbf{t})(\mathbf{e}^i \cdot \mathbf{t})\right)\mathbf{n}, \quad E_t \mathbf{e}^t \cdot \mathbf{n} \neq 0,$$
$$(4.48)$$

When $E_t = 0$, replace (4.48) with (4.44), and we may assign $\mathbf{e}^t$ arbitrarily, hence (4.47) is redundant. If $\mathbf{e}^t \cdot \mathbf{n} = 0$, then replace (4.48) with (4.44).

GO1 - energy flow argument

The time-averaged intensity of a time-harmonic wave associated with a solution u of the Helmholtz equation 4.15 is given by $\langle \mathbf{I} \rangle = C \mathrm{Im}\,\bar{u}\nabla u$, where C is a positive constant depending on the frequency (see, e.g., [60, §3.4.2]). When u is a plane wave of the form (4.17) this gives $\langle \mathbf{I} \rangle = C|A|^2 k_0 D \mathbf{d}$, so that the energy flow is purely in the real propagation direction $\mathbf{d}$.

We stipulate that the time-averaged intensities $\langle \mathbf{I}_i \rangle$ and $\langle \mathbf{I}_t \rangle$ associated with the incident and transmitted waves should satisfy

$$\mathrm{sgn}(\langle \mathbf{I}_i \rangle \cdot \mathbf{n}) = \mathrm{sgn}(\langle \mathbf{I}_t \rangle \cdot \mathbf{n}),$$

to ensure that the transmission process preserves the direction of energy flow relative to the boundary. By the above discussion this means that we require

$$\mathrm{sgn}(\mathbf{d}^t \cdot \mathbf{n}) = \mathrm{sgn}(\mathbf{d}^i \cdot \mathbf{n}),$$

so that (4.42) becomes, after applying Snell's law,

$$\mathbf{d}^t = \frac{D_i}{D_t}(\mathbf{d}^i \cdot \mathbf{t})\mathbf{t} + \mathrm{sgn}(\mathbf{d}^i \cdot \mathbf{n})\sqrt{1 - \left(\frac{D_i}{D_t}\right)^2 (\mathbf{d}^i \cdot \mathbf{t})^2}\,\mathbf{n}. \qquad (4.49)$$

This formula implies that the transmitted wave is always propagating into the second medium, except for the case of total internal reflection (TIR), i.e., when $\mathbf{d}^t \cdot \mathbf{n} = 0$ and there is no sign choice to be made. The energy flow in the second medium in this case is parallel to the interface.

We now turn to $\mathbf{e}^t$. It turns out that, having specified the sign choice in (4.42), the sign choice in (4.43) follows immediately. Indeed, provided that $E_t \mathbf{d}^t \cdot \mathbf{n} \neq 0$, $\mathbf{e}^t \cdot \mathbf{n}$ is now completely determined by (4.34), with

$$\mathbf{e}^t \cdot \mathbf{n} = \frac{1}{D_t E_t \mathbf{d}^t \cdot \mathbf{n}} \left(\mu_2 \xi_2 - D_i E_i (\mathbf{d}^i \cdot \mathbf{t})(\mathbf{e}^i \cdot \mathbf{t}) \right), \qquad E_t \mathbf{d}^t \cdot \mathbf{n} \neq 0.$$

Thus

$$\mathbf{e}^t = \frac{E_i}{E_t}(\mathbf{e}^i \cdot \mathbf{t})\mathbf{t} + \frac{1}{D_t E_t \mathbf{d}^t \cdot \mathbf{n}} \left(\mu_2 \xi_2 - D_i E_i (\mathbf{d}^i \cdot \mathbf{t})(\mathbf{e}^i \cdot \mathbf{t}) \right) \mathbf{n}, \qquad E_t \mathbf{d}^t \cdot \mathbf{n} \neq 0. \quad (4.50)$$

This equation predicts that the vector $\mathbf{e}^t$ sometimes points back into the first (incident) medium. This somewhat counterintuitive behaviour was noted in [116]; however, some subsequent authors (in particular, [32]), seemingly unaware of Pincherle's work, artificaly force $\mathbf{e}^t$ to point into the second medium, despite the fact that this may lead to a violation of the Helmholtz equation. We remark that a similar, artificial modification to the laws of reflection/refraction is made in [16] and [148], where the transmitted wave is spuriously forced to be homogeneous, when in practice it could be inhomogeneous as outlined above.

When $E_t = 0$ or $\mathbf{d}^t \cdot \mathbf{n} = 0$ the formula (4.50) cannot be applied. The former case is easily dealt with: since E_t and $\mathbf{e}^t$ appear in a product in the formula 4.17, the choice of $\mathbf{e}^t$ is irrelevant when $E_t = 0$, and we may arbitrarily assign $\mathbf{e}^t = \mathbf{d}^t$, for example. The latter case $\mathbf{d}^t \cdot \mathbf{n} = 0$ corresponds to TIR, and in this case we argue that the transmitted wave should decay (not grow) with increasing distance from the interface, so that we require $\mathrm{sgn}(\mathbf{e}^t \cdot \mathbf{n}) = \mathrm{sgn}(\mathbf{d}^i \cdot \mathbf{n})$, giving

$$\mathbf{e}^t = \frac{E_i}{E_t}(\mathbf{e}^i \cdot \mathbf{t})\mathbf{t} + \mathrm{sgn}(\mathbf{d}^i \cdot \mathbf{n})\sqrt{1 - \left(\frac{E_i}{E_t}\right)^2 (\mathbf{e}^i \cdot \mathbf{t})^2}\mathbf{n}, \qquad \mathbf{d}^t \cdot \mathbf{n} = 0. \quad (4.51)$$

GO2 - decay argument

In the zero absorption case, inhomogeneous waves arise when total internal reflection occurs. In this event, the transmitted wave propagates along the boundary and is attenuating perpendicular to this direction. There is a choice of two possible attenuation directions. The physics textbooks [19, 76, 130] choose $\mathbf{e}^t$ to be directed into the second medium to ensure that the wave decays as $y \to \infty$. Here, we shall make the same argument for the absorption case. To ensure this, we enforce the condition

$$\mathrm{sgn}(\mathbf{e}^t \cdot \mathbf{n}) = \mathrm{sgn}(\mathbf{d}^i \cdot \mathbf{n})$$

and hence equation (4.43) becomes, after applying Snell's Law,

$$\mathbf{e}^t = \frac{E_i}{E_t}(\mathbf{e}^i \cdot \mathbf{t})\mathbf{t} + \text{sgn}(\mathbf{d}^i \cdot \mathbf{n})\sqrt{1 - \left(\frac{E_i}{E_t}\right)^2 (\mathbf{e}^i \cdot \mathbf{t})^2}\mathbf{n}. \qquad (4.52)$$

Now we have specified $\mathbf{e}^t$, the direction $\mathbf{d}^t$ is determined uniquely by equation (4.34) as

$$\mathbf{d}^t = \frac{D_i}{D_t}(\mathbf{d}^i \cdot \mathbf{t})\mathbf{t} + \frac{1}{D_t E_t \mathbf{e}^t \cdot \mathbf{n}}\left(\mu_2 \xi_2 - D_i E_i (\mathbf{d}^i \cdot \mathbf{t})(\mathbf{e}^i \cdot \mathbf{t})\right)\mathbf{n}, \qquad E_t \mathbf{e}^t \cdot \mathbf{n} \neq 0. \ (4.53)$$

This argument leads to a propagation direction that may point back into the incident medium which seems counterintuitive, however we shall let experimental evidence dictate which is the correct direction to take.

When $E_t = 0$ or $\mathbf{e}^t \cdot \mathbf{n} = 0$ the formula (4.53) cannot be applied. The former case is dealt with as before, that is, since E_t and $\mathbf{e}^t$ appear in a product in the formula 4.17, the choice of $\mathbf{e}^t$ is irrelevant when $E_t = 0$, and we may arbitrarily assign $\mathbf{e}^t = \mathbf{d}^t$, for example. In the latter case $\mathbf{e}^t \cdot \mathbf{n} = 0$ we see that the vector $\mathbf{e}^t$ is parallel to the interface and therefore our decay argument cannot be applied. In this case, it makes sense to calculate $\mathbf{d}^t$ as in GO1, i.e.,

$$\mathbf{d}^t = \frac{D_i}{D_t}(\mathbf{d}^i \cdot \mathbf{t})\mathbf{t} + \text{sgn}(\mathbf{d}^i \cdot \mathbf{n})\sqrt{1 - \left(\frac{D_i}{D_t}\right)^2 (\mathbf{d}^i \cdot \mathbf{t})^2}\mathbf{n}. \qquad (4.54)$$

The Air/Ice and Ice/Air interface problems

As mentioned in the beginning of §4.3, we are interested mainly in the cases (ii) and (iv), that is,

- an interface between a non-absorbing medium and an absorbing medium - an "air/ice" interface (when light enters the crystal),

- an interface between an absorbing and non-absorbing medium - an "ice/air" interface (when light exits the crystal).

These are special cases of the general interface problem between two absorbing media considered until now. The application of the above derived formulae to these special cases is simple and so we do not provide versions of the formulae for these specific cases. However, we would like to comment on the applicability of GO1 and GO2 in these cases, and whether they ever coincide.

Air/Ice interface. At this interface, the upper medium is non-absorbing and the lower medium is absorbing. The incident wave is the incident homogeneous plane wave of our

scattering problem and hence $E_i = 0$. It can be seen from equations (4.45) and (4.48) that the vector $\mathbf{e}^t = \text{sgn}(\mathbf{d}^i \cdot \mathbf{n})\mathbf{n}$ and so for this case GO1 and GO2 coincide.

We remark that this case is covered in textbooks such as [19] and that their given formulae agree with those derived here.

Ice/Air interface. At this interface, the upper medium is absorbing and the lower medium is non-absorbing. The incident wave is an inhomogeneous plane wave and hence $E_i \neq 0$. For this interface, GO1 and GO2 are always distinct with one exception: when $\mathbf{e}^i \cdot \mathbf{t} = 0$. When $\mathbf{e}^i \cdot \mathbf{t} = 0$ it can be shown from (4.41) that $E_t = 0$ and so the choice of $\mathbf{e}^t$ is irrelevant and we choose $\mathbf{d}^t$ to point into the lower medium.

We remark that this case has only been covered in textbooks when $\mathbf{e}^i \cdot \mathbf{t} = 0$ (which arises when analysing absorbing layers, see [19]) and their given formulae agree with these derived here for this case.

4.3.4   Sign choice experiment

We have discussed two plausible physical arguments for specifying $\mathbf{d}^t$ and $\mathbf{e}^t$. In this section we present the results of a numerical experiment in which we compare the boundary data u_{go} and $\frac{\partial u_{go}}{\partial \mathbf{n}}$ computed using the two different choices GO1 and GO2 with the "exact" solution computed using the conventional BEM outlined in §3. We consider scattering by the quadrilateral shown in Figure 4.3 in which sides Γ_1 and Γ_4 have length 2π, Γ_2 has length π and Γ_3 has length $\sqrt{3}\pi$. We shall consider 100 different incident plane waves with incident angles θ^i spaced evenly between $\pi/6$ and $\pi/2$, inclusive, as shown in the figure. We also take $\alpha = 1$ throughout.

This particular shape has been chosen to isolate some behaviour from the equilateral triangle example considered in §3.2.4, which shall be the main example in §5. We see that sides Γ_1 and Γ_4 are related to each other in the same way as Γ_1 and Γ_3 are in the triangle case (see Figure 3.4). The other sides have been chosen such that the beams refracted through them do not impinge on the red portion of Γ_4, which we shall denote Γ_R. Our aim is to isolate the single beam which refracts through Γ_1 and then through Γ_R. To help achieve this we choose the imaginary part of the refractive index large enough so that the interior waves decay sufficiently rapidly to make re-reflections negligibly small. In particular, we choose n_I such that $k_1 n_I = 1$ so that the largest amplitude of an interior reflected beam impinging on Γ_R is 1.1×10^{-3} which is approximately 1000 times smaller than the magnitude of the aforementioned single beam we have endeavoured to isolate. We are here interested in the beam propagating out of Γ_R since this represents the "difficult" interface problem of absorbing to non-absorbing.
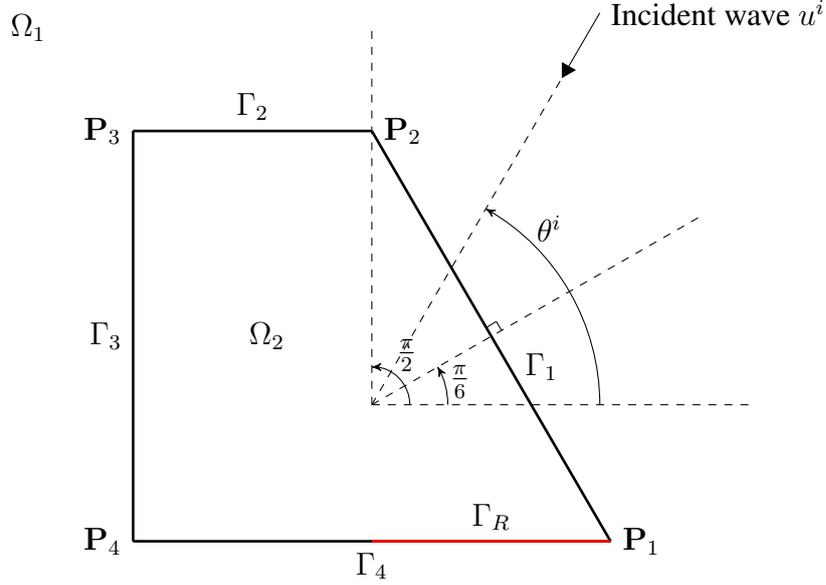
Figure 4.3: Setup used in numerical experiments.

We begin by taking the frequency of the incident wave to be $k_1 = 20$, so that $n_I = 0.05$. The real part of the refractive index we shall take throughout to be $n_R = 1.5$. In Figure 4.4 we present the relative L^2 errors for u_{go} and $\frac{\partial u_{go}}{\partial \mathbf{n}}$ on Γ_R, as well as the dot product of the transmitted vector $\mathbf{d}^t$ with the tangent vector to Γ_R, and its derivative with respect to the incident angle θ^i.
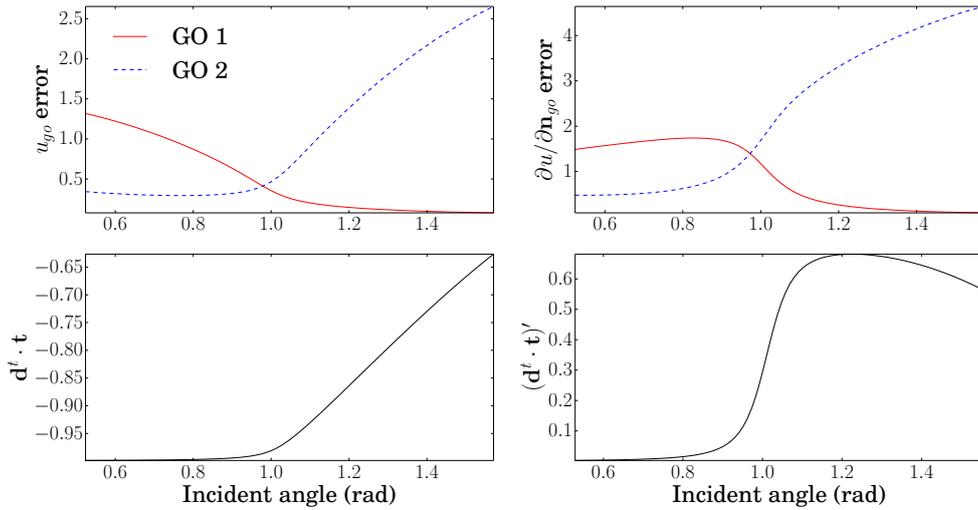
Figure 4.4: Top two plots show a comparison of the relative L^2 errors in u and $\partial u/\partial \mathbf{n}$ on the red segment Γ_R for the two different beam tracing methods. Bottom two plots display $\mathbf{d}^t \cdot \mathbf{t}$ and its derivative with respect to θ^i. Problem parameters: $k_1 = 20, n = 1.5 + 0.05\mathrm{i}$.

There are two distinct regimes to be seen in the plots. The first is located at $\theta^i < 0.98$ rad where GO2 provides the more accurate approximation. In this θ^i region we notice that $\mathbf{d}^t \cdot \mathbf{t}$ is very close to -1, i.e., the wave is close to total internal reflection, and its θ^i derivative is close to 0. More precisely, we have that $\mathbf{d}^t \cdot \mathbf{t} < -0.99$. The second regime is located at $\theta^i > 0.98$ rad where GO1 provides the more accurate approximation. In this region, $\mathbf{d} \cdot \mathbf{t}$ is moving away from -1 and towards 0 at an approximately linear rate. It should be noted that in a neighbourhood of this cross-over point, the GO approximation for $\partial u / \partial \mathbf{n}$ has approximately 100% error with either sign choice, indicating that in this range of θ^i, GO appears to be completely invalid.

Let us repeat the experiment but this time with larger k_1. We shall consider $k_1 = 40$ and $k_1 = 80$ in order to see how the behaviour exhibited in Figure 4.4 changes as the frequency increases. In order to maintain the same strength of the field on Γ_R, we shall decrease the imaginary component of the refractive index in order to maintain the relationship $k_1 n_I = 1$. That is, for $k_1 = 40$ we have $n = 1.5 + 0.025$i and for $k_1 = 80$, $n = 1.5 + 0.0125$i.

The corresponding versions of Figure 4.4 are shown below in Figure 4.7 and Figure 4.8. Similar behaviour is noted with the addition that the neighbourhood of the cross-over point in which GO appears invalid becomes narrower as k_1 increases. We also observe that the cross-over point shifts slightly towards larger θ^i, but this is merely due to the change of the imaginary part of the refractive index. This affects the effective refractive index, with real and imaginary parts denoted D and E in §4.3.2 which are functions of n_R and n_I, and hence the critical angle (corresponding to TIR) will be altered. Finally, we note the accuracy of the GO approximation improving as a whole from Figure 4.4 through Figure 4.5 to Figure 4.6, as is to be expected from an asymptotic approximation valid in the limit $k_1 \to \infty$.

To further illustrate the behaviour of the two GO approximations in the different θ^i regions, we present two figures below. They show the exact solution and the GO approximation on the red portion of Γ_4 for a value of θ^i in each of the three different regions. We see in Figure 4.7 that GO 1 is out of phase in the first region, approximates the amplitude poorly in the transition region and performs excellently in approximating the phase and amplitude in the final region. To contrast, we see in Figure 4.8 that GO 1 approximates the phase and amplitude well in the first region, matches the phase but approximates the amplitude with some error in the transition region, and approximates the amplitude as far too large in the final region. These observations are consistent with what we saw earlier, namely that, GO 1 is accurate in the final region, GO 2 is accurate in the first region, and both appear to perform badly near the cross-over point at $\mathbf{d}^t \cdot \mathbf{t} \approx -0.99$.

On the basis of these numerical experiments we can propose a simple rule of thumb to decide when to switch from one sign choice to the other. This would ideally be at the
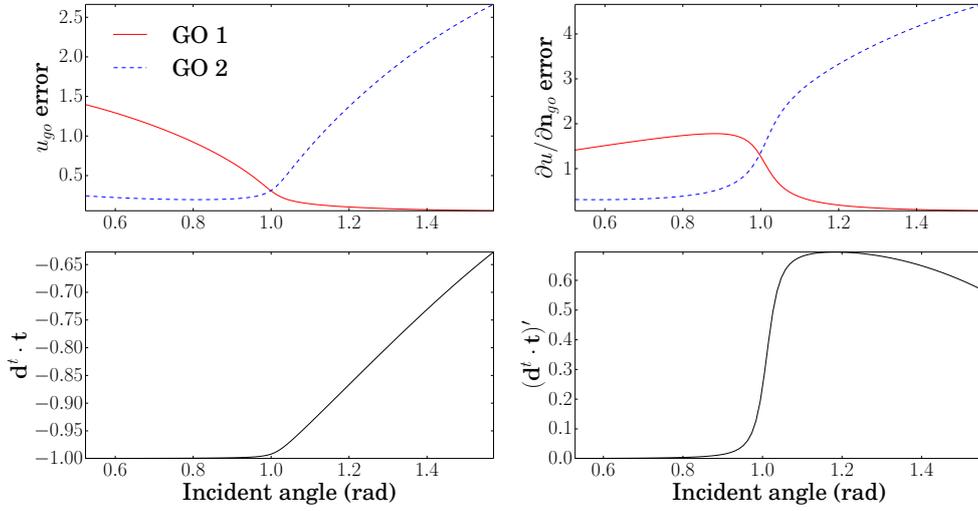
Figure 4.5: As Figure 4.4 but with $k_1 = 40, \mu = 1.5 + 0.025\text{i}$.
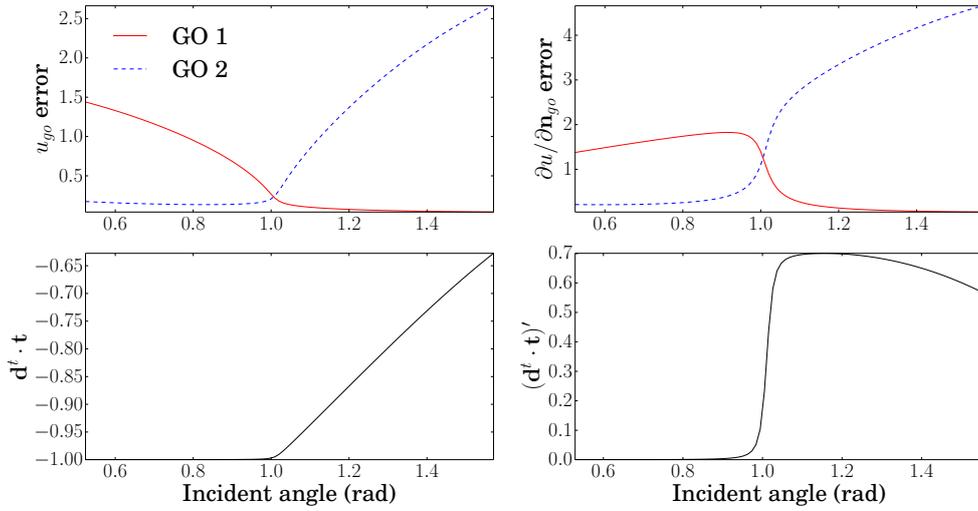
Figure 4.6: As Figure 4.4 but with $k_1 = 80, \mu = 1.5 + 0.0125\text{i}$.

observed cross-over point. The value of $\mathbf{d}^t \cdot \mathbf{t}$ at this point is typically between -1 and -0.98. A sensible ad hoc rule that we may establish for switching could be: *if $\mathbf{d}^t \cdot \mathbf{t} < -0.99$, use GO 2, else use GO 1*. Or, more generally,

$$\textit{if } |\mathbf{d}^t \cdot \mathbf{t}| > 0.99\textit{, use GO 2, else use GO 1}. \tag{4.55}$$
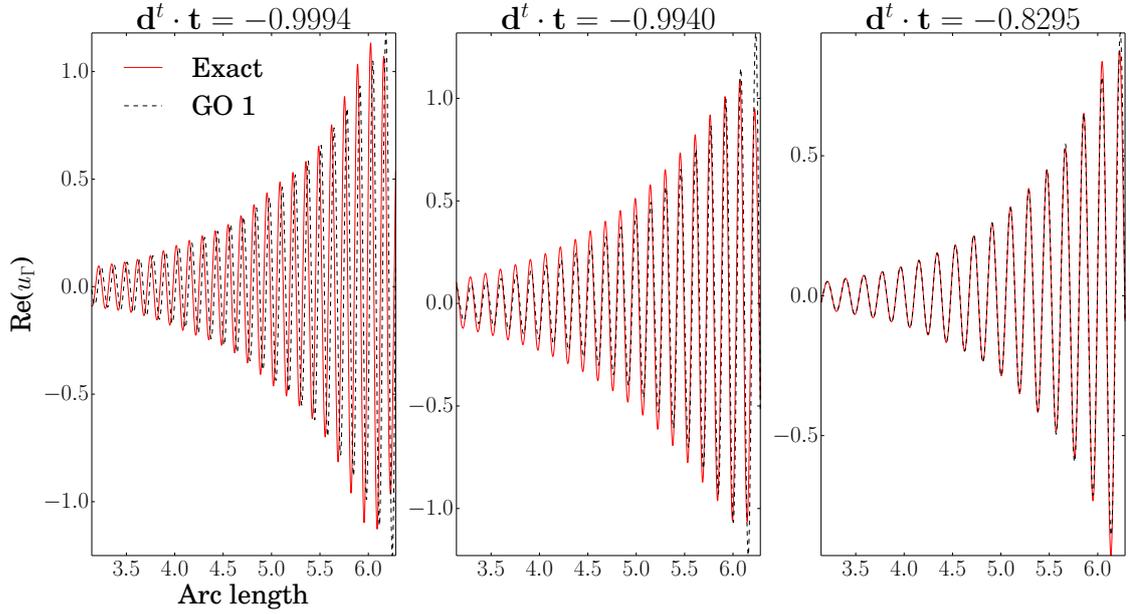
Figure 4.7: $\mathrm{Re(u)}$ on Γ_R compared with the GO1 approximation for three different values of θ_i corresponding to $\mathbf{d}^t \cdot \mathbf{t} = -0.9994, -0.9940, -0.8295$. Problem parameters: $k_1 = 40$ and $n = 1.5 + 0.025\mathrm{i}$.
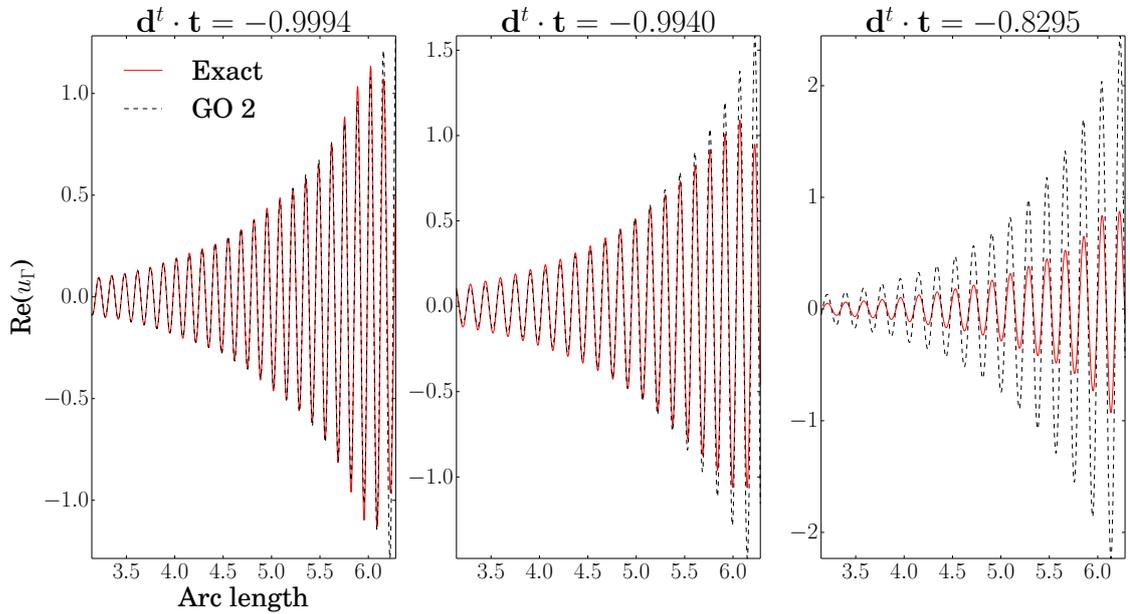
Figure 4.8: $\mathrm{Re(u)}$ on Γ_R compared with the GO2 approximation for three different values of θ_i corresponding to $\mathbf{d}^t \cdot \mathbf{t} = -0.9994, -0.9940, -0.8295$. Problem parameters: $k_1 = 40$ and $n = 1.5 + 0.025\mathrm{i}$.

4.4 Convergence of beam tracing algorithm

As depicted in Figure 4.1 and described in §4.2, the GO approximation as constructed via a BTA is composed of numerical beams arising from internal reflections of the "primary" transmitted beams. By primary, we refer to the beams transmitted into the shape as the incident plane wave strikes the illuminated sides. Suppose we were to terminate the beam tracing algorithm here. Then the GO field would be given by

$$\mathbf{v}_{go}(\mathbf{x}) = \sum_{l=1}^{N_1} \left(u_{1l}^t(\mathbf{x}, k_2), \alpha \frac{\partial u_{1l}^t}{\partial \mathbf{n}}(\mathbf{x}, k_2) \right), \quad \mathbf{x} \in \Gamma, \tag{4.56}$$

where N_1 is the number of sides illuminated by the incident wave, and hence the number of beams. For each beam, the superscript t denotes the transmitted field, and the subscript 1 signifies that these are the primary beams. Owing to the transmission conditions, we could alternatively write this field as

$$\mathbf{v}_{go}(\mathbf{x}) = \left(u^i(\mathbf{x}, k_1), \frac{\partial u^i}{\partial \mathbf{n}}(x, k_1) \right) + \sum_{l=1}^{N_1} \left(u_{1l}^r(\mathbf{x}, k_1), \frac{\partial u_{1l}^r}{\partial \mathbf{n}}(\mathbf{x}, k_1) \right), \quad \mathbf{x} \in \Gamma, \tag{4.57}$$

where the superscript denotes the corresponding reflected beams. Also, we note that $\mathbf{v}^i = \left(u^i(\mathbf{x}, k_1), \frac{\partial u^i}{\partial \mathbf{n}}(x, k_1) \right)$ is only supported on the illuminated sides. We shall choose the representation (4.56) for sake of brevity.

Now suppose that we continue to track the transmitted beams to their next reflection/refraction, and then calculate the GO field. Then we have

$$\mathbf{v}_{go}(\mathbf{x}) = \sum_{l=1}^{N_1} \left(u_{1l}^t(\mathbf{x}, k_2), \alpha \frac{\partial u_{1l}^t}{\partial \mathbf{n}}(\mathbf{x}, k_2) \right) + \sum_{l=1}^{N_2} \left(u_{2l}^t(\mathbf{x}, k_1), \frac{\partial u_{2l}^t}{\partial \mathbf{n}}(\mathbf{x}, k_1) \right), \tag{4.58}$$

where N_2 is the number of "secondary" transmitted beams, and $(u_{2l}^t, \partial u_{2l}^t / \partial \mathbf{n})$, $l = 1, \dots, N_2$ correspond to the beams transmitted out of the shape at this reflection/refraction event. There are also reflected beams generated which propagate back into the shape leading to higher order reflected/refracted beams.

Say we consider M reflections in total in the calculation of the GO field, so that

$$\mathbf{v}_{go}(\mathbf{x}) = \sum_{l=1}^{N_1} \left(u_{1l}^t(\mathbf{x}, k_2), \alpha \frac{\partial u_{1l}^t}{\partial \mathbf{n}}(\mathbf{x}, k_2) \right) + \sum_{m=2}^{M} \sum_{l=1}^{N_m} \left(u_{ml}^t(\mathbf{x}, k_1), \frac{\partial u_{ml}^t}{\partial \mathbf{n}}(\mathbf{x}, k_1) \right). \tag{4.59}$$

A natural question to ask is how large a value of M should we take? This is the topic of this section. We would like to determine the number of reflections M required in our algorithm for the solution to converge to a given tolerance.

This question was analysed in [52] for the same beam tracing algorithm as that employed here, however there the scatterers considered were all non-absorbing. It was shown that as many as 40 reflections were sometimes required (for some non-absorbing scatterers) before the relative error in the BTA reduced to approximately machine precision. In that study, as we shall do here, the reference boundary solution U_{ref} was taken as that calculated by the BTA with $M = 50$. The error studied was the relative L^2 error. That is, the error in the solution u_M, which is that calculated by the BTA with $M = m$, was defined as

$$e(M) = \frac{||U_{ref} - u_M||_{L^2(\Gamma)}}{||U_{ref}||_{L^2(\Gamma)}}. \tag{4.60}$$

We shall also use this error in the study here. However, since the error of the GO approximation is typically a few percent (as we shall see later in this section, and in §4.5), it is not necessary to take M large enough to ensure $e(M) \approx \varepsilon_{machine}$. In fact, it is beneficial to have M as small as possible for when we employ the HNA algorithm later since this will reduce the number of integrals to be evaluated (see §6). Therefore, here we shall have a different aim to that in [52]. That is, the aim of obtaining a practical rule for truncating the series (4.59) at as small an M as possible without allowing the accuracy of the GO to be compromised.

In order to ascertain such a rule, we consider some examples in which we calculate M to achieve two different tolerances for $e(M)$: machine precision (as was done in [52]) and our specified tolerance of $tol = 10^{-3}$. The first tolerance value will lead us to the value of M at which taking any further terms would have no effect upon the solution. We denote this M_ε. The second tolerance value was chosen in order to produce a smaller M, which we denote M_{tol}, more practical for computations. The value of $tol = 10^{-3}$ is chosen because the relative error of the GO approximation is never below 3×10^{-2} for any of our examples (as we shall see in §4.5 and §5), so including beams with amplitudes smaller than this tolerance will have an almost negligible effect on the accuracy of the GO approximation.

For each example, we shall also consider the amplitudes of the beams of orders M_{tol} and $M_{tol} + 1$. Denote the amplitude of each beam of order M_{tol} by $A_{M_{tol},l}$ for $l = 1, \ldots, N_{M_{tol}}$. Then we shall record the maximum of these amplitudes and denote it as $||A_{M_{tol}}||_\infty$. We shall also do the same for $A_{M_{tol}+1,l}$. This will allow us to derive a rule to terminate the BTA by analysing the amplitudes of the beams. Finally, for each example, we provide the relative error in the GO approximation (compared against an exact solution computed using the 2D BEM of §3) calculated with $M = M_\varepsilon$ and $M = M_{tol}$ to demonstrate that little or no accuracy is lost in the GO approximation by truncating at $M = M_{tol}$ rather than at $M = M_\varepsilon$.

90

The first example we consider is scattering by the triangle in Figure 3.4 with an incident direction specified by Angle 2. We choose three different refractive indices for this example: $n = 1.5 + 0i$, $n = 1.5 + 0.025i$ and $n = 1.31 + 0i$. The results are shown in Tables 4.9, 4.10 and 4.11. We can immediately see from the tables that M_{tol} is always considerably

k_1	M_ε	$\frac{\|u-u_{go}^{M_\varepsilon}\|}{\|u\|}$	M_{tol}	$\frac{\|u-u_{go}^{M_{tol}}\|}{\|u\|}$	$\|A_{M_{tol}}\|_\infty$	$\|A_{M_{tol}+1}\|_\infty$
10	36	2.49e-1	7	2.49e-1	7.02e-3	2.64e-3
40	36	1.62e-1	7	1.62e-1	7.02e-3	2.64e-3
160	36	1.10e-1	7	1.10e-1	7.02e-3	2.64e-3

Table 4.9: Convergence details for the BTA for scattering of a plane wave at Angle 2 by a triangle with $n = 1.5 + 0i$.

k_1	M_ε	$\frac{\|u-u_{go}^{M_\varepsilon}\|}{\|u\|}$	M_{tol}	$\frac{\|u-u_{go}^{M_{tol}}\|}{\|u\|}$	$\|A_{M_{tol}}\|_\infty$	$\|A_{M_{tol}+1}\|_\infty$
10	21	1.70e-1	5	1.70e-1	1.20e-2	1.20e-3
40	12	9.11e-2	3	9.11e-2	3.98e-2	5.21e-4
160	5	4.77e-2	2	4.86e-2	7.38e-1	3.98e-2

Table 4.10: As per Table 4.9 but with $n = 1.5 + 0.025i$.

k_1	M_ε	$\frac{\|u-u_{go}^{M_\varepsilon}\|}{\|u\|}$	M_{tol}	$\frac{\|u-u_{go}^{M_{tol}}\|}{\|u\|}$	$\|A_{M_{tol}}\|_\infty$	$\|A_{M_{tol}+1}\|_\infty$
10	28	2.15e-1	6	2.15e-1	6.09e-3	1.08e-3
40	28	1.51e-1	6	1.51e-1	6.09e-3	1.08e-3
160	28	1.04e-1	6	1.04e-1	6.09e-3	1.08e-3

Table 4.11: As per Table 4.9 with $n = 1.31 + 0i$.

smaller than M_ε and that only in one case ($k_1 = 160$ in Table 4.10) is the accuracy of u_{go} affected (to three significant figures) by taking $M = M_{tol}$. In this case, the accuracy of u_{go} diminishes by 0.11% which is small when compared to the accuracy of u_{go} itself, which is approximately 5%.

From looking at the values in the $\|A_{M_{tol}}\|_\infty$ and $\|A_{M_{tol}+1}\|_\infty$ columns, we can see that the maximum amplitude of the order M beams is always greater than 5×10^{-3} and that the maximum amplitude of the order $M + 1$ beams which is often less than 5×10^{-3} (note there is only one instance in the examples considered in tables 4.9–4.11 where this value exceeds 5×10^{-3}, namely for $n = 1.5 + 0.025i$ and $k_1 = 160$). Therefore, it appears that, to ensure $e(M) < tol$, we can terminate the BTA at M once the amplitudes of all the beams of the $(M + 1)$st iteration have a magnitude less than 5×10^{-3} (for the exceptional case,

this will lead to us including quite a few more beams than necessary which will lead to a higher computational expense but will not deteriorate the accuracy).

Based on the above, we shall impose the following truncation rule in our BTA:

In order to achieve
$$e(m) < 10^{-3}$$
we shall terminate the algorithm at $m = M$ when
$$||A_{M+1}||_\infty < 5 \times 10^{-3}. \qquad (4.61)$$

To reiterate this rule in words: in order to achieve our imposed tolerance of $e(m) < 10^{-3}$, at each iteration M we check the value of A_{M+1} (the amplitudes of the beams from the next iteration) for every beam of order $M + 1$. If *every* beam of order $M + 1$ satisfies the criterion $|A_{M+1}| < 5 \times 10^{-3}$, then we terminate the BTA at M iterations, if not then we proceed to the next iteration.

It may seem surprising that we can include beams of amplitude larger than 10^{-3} in order to achieve $e(m) < 10^{-3}$. However this apparent inconsistency is due to the fact that the support of each beam is substantially smaller than the whole of Γ. Therefore, the magnitude of the contribution of each beam to the total u_{go} approximation is proportional to that beam's amplitude multiplied by the ratio of the length of its support to the length of Γ.

In this section we have presented some results illustrating the accuracy of the GO approximation on the boundary of the scatterer for a small selection of examples. In the following section we shall provide more complete results displaying how the accuracy changes with increasing k_1 and increasing absorption. We also provide results illustrating the accuracy of the Kirchhoff approximation (based on this GO approximation) in the domain and far-field.

4.5   Accuracy of the GO approximation

The accuracy of high-frequency asymptotic approximations for wave problems typically improve with increasing k_1 with a cost fixed independently of k_1. For a fixed k_1, however, it is not in general possible to improve the accuracy of the approximation to any desired tolerance. In this section, we shall examine the accuracy of the GO approximation computed by the BTA developed in previous sections. The BTA shall be applied in detail to the triangular scatterer in Figure 3.4 at five different levels of absorption and for incident waves of wavenumber $k_1 = 10, 20, 40, 80, 160$. We shall see that as k_1 increases, the accuracy of

the GO approximation improves. However, the GO approximation on the boundary is only ever (for the range of k_1 considered) accurate to at best a few percent, and for low k_1, the accuracy is poor (as is to be expected).

Upon substituting the GO approximation for the boundary data $(u, \partial u/\partial \mathbf{n})$ into the Green's representation formula (2.78) or its far-field version (2.127), we obtain the so-called "Kirchhoff approximation" (KA) or "Physical Optics approximation" . This is a commonly used technique for simulating wave scattering by ice crystals as well as other applications [108, 147] owing to its relative simplicity and low computational cost. As we saw in §3.2.4, it is to be expected that the relative error is smaller in the domain and far-field than on the boundary. We also present results of the KA for the aforementioned examples.
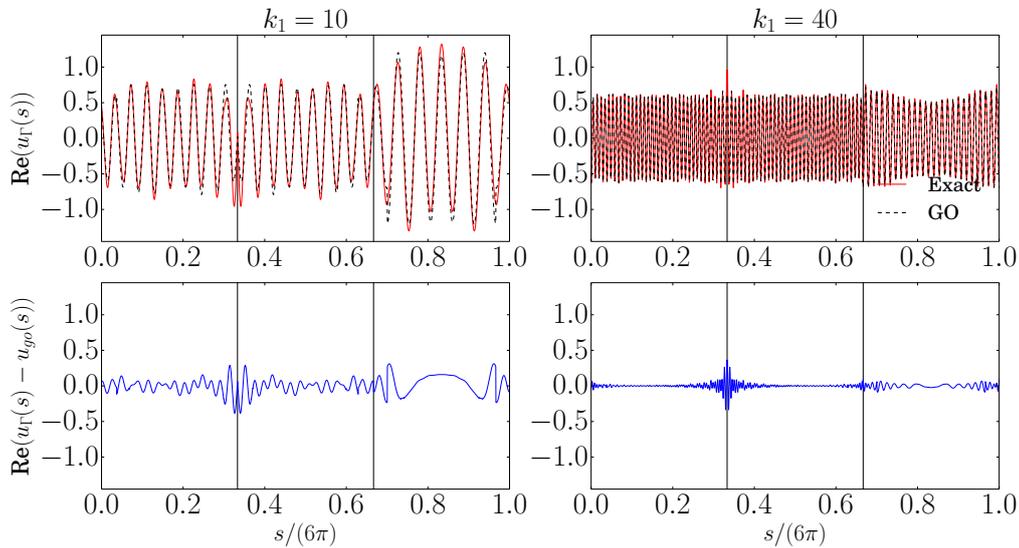
Figure 4.12: Boundary solution compared to GO approximation for scattering setup 2 (inc $\pi/2$) of Figure 3.4 with refractive index $1.5 + 0.00625i$.

The results presented shall demonstrate that diffraction accounts for a significant portion of the field and must be taken into account if accurate approximations are required. The sections to follow are dedicated to developing an understanding of the diffracted field.

Presently, we consider scattering by the triangle mentioned above. We fix the real part of the refractive index at $n_R = 1.5$ and vary n_I to investigate the effect absorption has on the accuracy of the GO approximation. Figure 4.12 shows the real part of u and its GO approximation on the boundary for this example with $n_I = 1/160$. Also plotted is the difference between u and u_{go} which we identify with the diffracted field. At first glance we see that for $k_1 = 10$ the GO approximation reproduces qualitatively the correct solution, however further inspection of the difference $u - u_{go}$ reveals that there is a significant diffracted field which remains to be approximated. As k_1 is increased (to $k_1 = 40$ in the

figure), it is apparent that the GO solution improves significantly with the diffracted field becoming ever more confined close to the corners.

Figure 4.13 shows the relative L^2 errors of u_{go} and $\partial u_{go}/\partial \mathbf{n}$ on the boundary of the scatterer. It is clear that for all values of n_I considered, the accuracy of the GO approximation improves with increasing k_1. Also, for scatterers possessing absorption ($n_I > 0$), the GO approximation is better than that for the scatterer with no absorption. This is to be expected since for absorbing scatterers, the waves propagating within the shape are exponentially decaying and so the diffracted waves propagating within the scatterer are considerably weaker. However, upon more detailed inspection, Figure 4.13 shows that it is not necessarily the case that, the stronger the absorption, the better the GO approximation. This is explained by the fact that, as $k_1 n_I$ increases, the transmission problem becomes increasingly better approximated by an impenetrable scattering problem. For such a problem, the solution is in fact singular at the corners. Therefore, as $k_1 n_I$ increases, the solution becomes increasingly peaked at the corners. This corner behaviour is not captured by the GO and hence its accuracy deteriorates as $k_1 n_I$ increases.
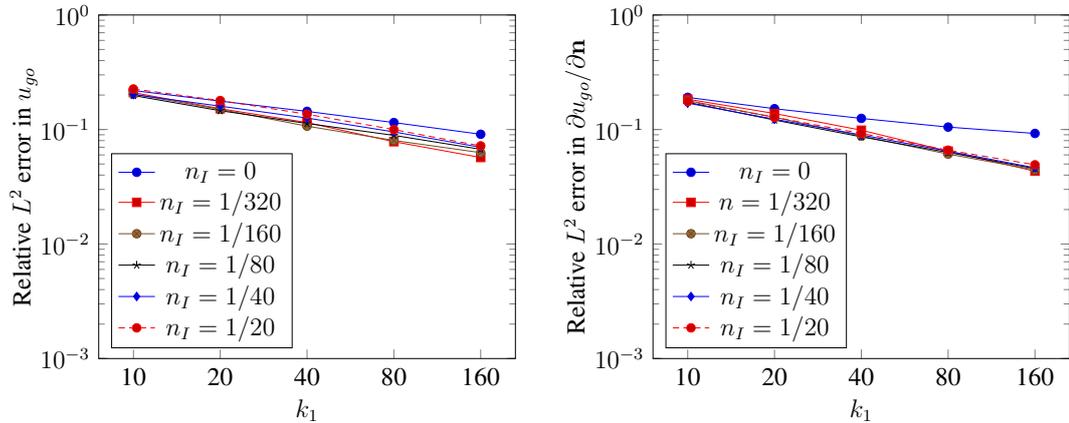
Figure 4.13: Accuracy of the GO approximation to u and $\partial u/\partial \mathbf{n}$ on Γ for scattering of an incident wave with direction $\mathbf{d}^i = (\cos(\pi/6), -\sin(\pi/6))$ by the triangle in Figure 3.4 with $n = 1.5 + n_I \mathrm{i}$.

4.6 Kirchhoff approximation

Now we shall consider the "Kirchhoff approximation" (KA), obtained by substituting the GO solution on the boundary into Green's representation formulae (2.74)–(2.75) or their far-field form. We briefly study the accuracy of the Kirchhoff approximation for the example considered in §4.5. Figure 4.14 shows the absolute value of the total field on a circle

of radius $3\pi/\sqrt{3}$ (3/2 times the radius of smallest circle which enscribes the triangle) surrounding the triangle, with $t/(2\pi) = 0$ corresponding to the positive x-direction. The red line is the exact solution and the dashed black line is the Kirchhoff approximation. For $k_1 = 10$ the KA performs well and reproduces the key features of the field, particularly in the direct forward scattering direction and the two regions corresponding to specular reflection from the sides Γ_1 and Γ_2. For $k_1 = 40$ these specular reflection regions, as well as the region in shadow, are even more distinct, and we notice that the KA approximation is virtually indistinguishable from the exact solution to the naked eye. However, as Figure 4.15 shows, the relative L^2 error of the KA on this circle is still quite large: approximately 10% for $k_1 = 10$ and 5% for $k_1 = 40$.
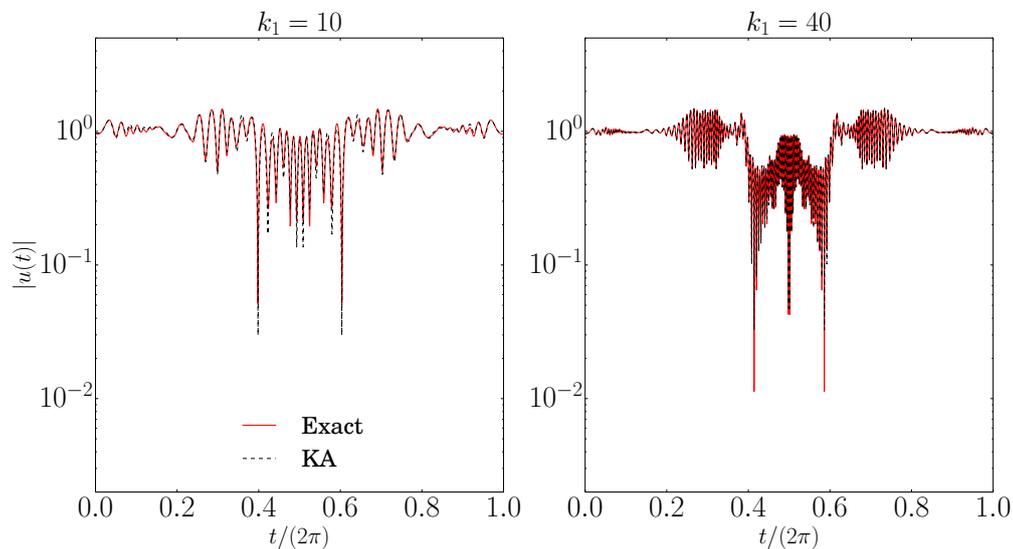
Figure 4.14: Total field (u) on a circle of radius $3/2\pi$ surrounding the triangular scatterer in Figure 3.4 with $\mu = 1.5 + \frac{1}{160}\mathrm{i}$. The incident wave direction is $\mathbf{d}^i = (0, -1)$.

Finally, in Figure 4.16 we display the errors of the KA in the far-field for the same scattering problem as was the subject of Figure 4.13. Here we see a similar pattern to the boundary errors displayed in Figure 4.13, however, in this case we have that the greater the absorption, the better the accuracy. Also, as is to be expected from an asymptotic approximation, the error decreases with increasing k_1. We see that accuracy is never better than 3% for all the levels of absorption and wavenumbers considered.

The results demonstrate that while the largest contribution to the scattered field comes from the geometrical optics, the remainder, which we identify with the diffracted field, also makes a significant contribution. It is the accurate approximation of this diffracted field which we shall consider in the remainder of the thesis. The goal is to determine the phases associated with the diffracted field in order to derive an appropriate hybrid
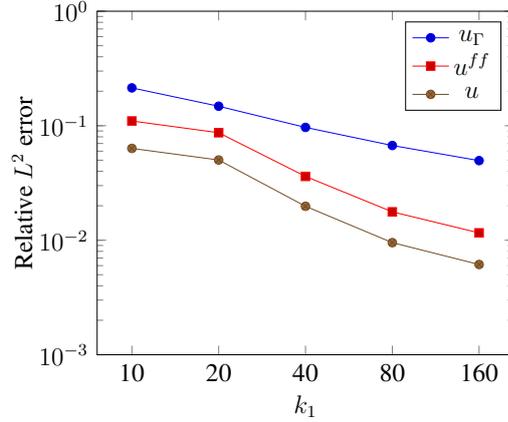
Figure 4.15: Relative errors of u_{go} on Γ, and the Kirchhoff approximation to the total field in the domain (u) and in the far field (u^{ff}) against k_1 for scattering by the triangle in Figure 3.4 with $n = 1.5 + \frac{1}{160}$i. The incident wave has direction $\mathbf{d}^i = (0, -1)$.
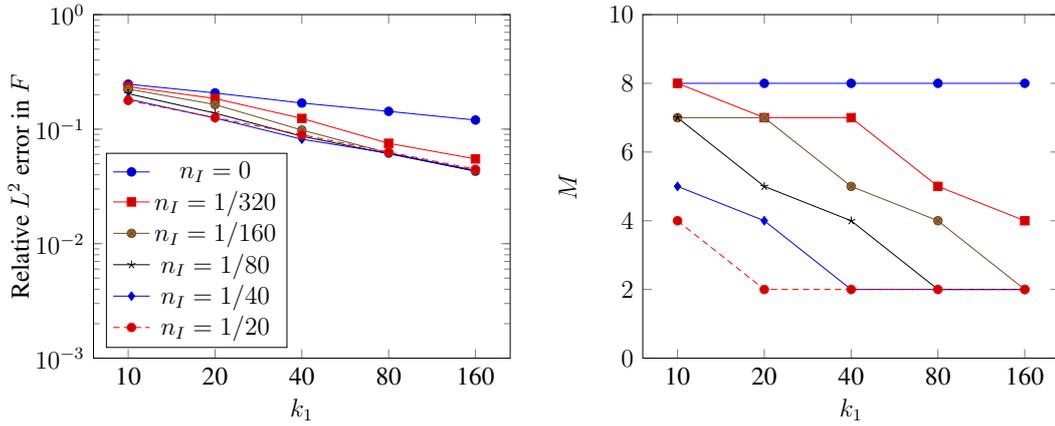
Figure 4.16: Left: Relative L^2 errors of the KA in the far-field for scattering by the triangle in Figure 3.4 with refractive index $n = 1.5 + n_I$i. The incident wave has direction $\mathbf{d}^i = (\cos(\pi/6), -\sin(\pi/6))$. Right: Number M of reflections in the GO approximation.

numerical-asymptotic approximation space for the transmission problem. To do so, we first review the basic ideas of the GTD.

4.7 Geometrical Theory of Diffraction

The Geometrical Theory of Diffraction was developed by Joseph Keller [79] as a high-frequency approach to approximate diffracted wave-fields. His theory was applicable to 2D and 3D problems of diffraction by corners, vertices, edges and smooth surfaces of positive curvature. Here we shall concentrate on those aspects of the GTD which are relevant to 2D diffraction by corners since we are interested in polygonal scatterers. For a more detailed

exposition of the GTD see, e.g., [20, 74, 79].

As we saw in earlier sections, at high frequencies the scattered field is often well-approximated by the reflected and refracted rays of geometrical optics. However, these rays do not include those arising from the phenomenon of diffraction. The main postulate of the GTD is that diffracted fields are generated by the rays from the incident field which strike the corners of the scattering surface. More precisely, when a ray from the incident field strikes a corner, infinitely many diffracted waves are generated travelling in all directions away from that corner, i.e., a cylindrical wave is created.

In this section we shall review some of the material for 2D GTD for impenetrable scatterers for which it has been well developed. Later we shall heuristically extend the principles of GTD for impenetrable scatterers to the penetrable case for use in the problem of interest in this thesis.

The GTD postulates that the diffracted field generated by a corner is related to the incident field as

$$u_d = \frac{e^{ikr}}{\sqrt{kr}} D(\theta, \theta_0, k) u^i, \tag{4.62}$$

where θ_0 is the angle of incidence of a plane wave; (r, θ) are the polar coordinates of the observation point; and D is the diffraction coefficient. The diffraction coefficient is calculated from the exact solution of the canonical problem of scattering by an infinite wedge which we shall briefly review here.

4.7.1 Sommerfeld problem

Consider an incident plane wave $u^i = e^{-ikr\cos(\theta-\alpha)}$ impinging on the semi-infinite line $y < 0$ where (r, θ) are polar coordinates centred at the origin. We impose the Dirichlet boundary condition on the half-line. Therefore we wish to solve

$$(\nabla^2 + k^2)u = 0, \tag{4.63}$$

$$u = 0, \quad x > 0, \ y = 0, \tag{4.64}$$

along with a suitable outgoing radiation condition for the scattered field.

This problem was first solved by Sommerfeld in 1896 using a generalisation of the method of images to a two-sheeted Riemann surface [126]. Other possible solution methods include the Wiener-Hopf technique (see, e.g., [17]) and the Kontorovich-Lebedev transform (see, e.g., [113]). All of these approaches lead to the classical result

$$u = \frac{e^{ikr-i\pi/4}}{\sqrt{\pi}} \left(-\mathrm{Fr}\left(\sqrt{2kr}\cos\left(\frac{\theta-\alpha}{2} \right) \right) + \mathrm{Fr}\left(-\sqrt{2kr}\cos\left(\frac{\theta+\alpha}{2} \right) \right) \right), \tag{4.65}$$
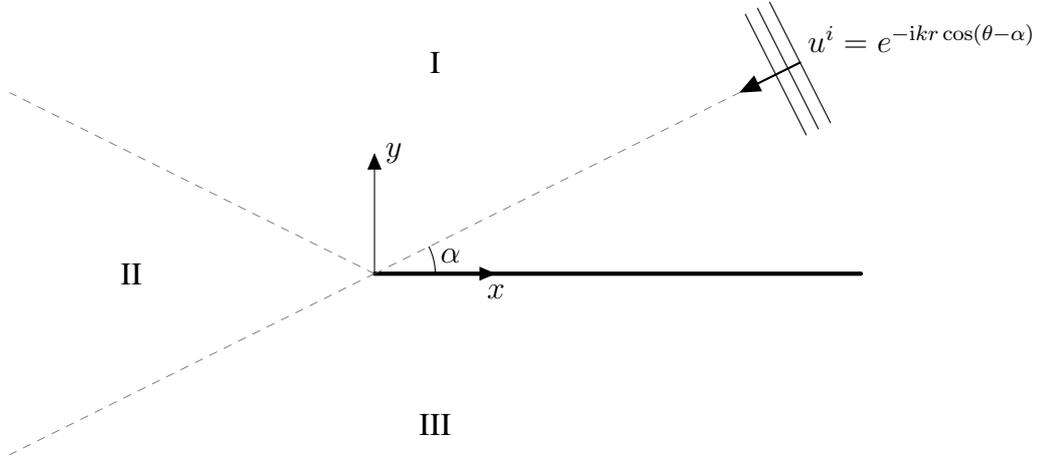
Figure 4.17: Scattering of a plane wave by a sound soft half line, also known as the *Sommerfeld problem*.

where

$$\mathrm{Fr}(z) = \mathrm{e}^{-\mathrm{i}z^2} \int_z^\infty \mathrm{e}^{\mathrm{i}s^2}\mathrm{d}s \tag{4.66}$$

is the Fresnel integral.

The behaviour of the solution for $kr \gg 1$ can be obtained via the large argument expansions of the Fresnel integral:

$$\mathrm{Fr}(z) \sim \frac{\mathrm{i}}{2z} + \mathcal{O}\left(\frac{1}{|z|^2}\right), \quad \mathrm{Re}(z) > 0, \tag{4.67}$$

$$\mathrm{Fr}(z) \sim \sqrt{\pi}\mathrm{e}^{\mathrm{i}\pi/4}\mathrm{e}^{-\mathrm{i}z^2} + \frac{\mathrm{i}}{2z} + \mathcal{O}\left(\frac{1}{|z|^2}\right), \quad \mathrm{Re}(z) < 0, \tag{4.68}$$

giving

$$u \sim \begin{cases} \mathrm{e}^{-\mathrm{i}kr\cos(\theta-\alpha)} + \mathrm{e}^{-\mathrm{i}kr\cos(\theta+\alpha)} + D(\theta)\frac{\mathrm{e}^{\mathrm{i}kr}}{\sqrt{kr}}, & 0 < \theta < \pi - \alpha, \\ \mathrm{e}^{\mathrm{i}kr\cos(\theta-\alpha)} + D(\theta)\frac{\mathrm{e}^{\mathrm{i}kr}}{\sqrt{kr}}, & \pi - \alpha < \theta < \pi + \alpha, \\ D(\theta)\frac{\mathrm{e}^{\mathrm{i}kr}}{\sqrt{kr}}, & \pi + \alpha < \theta < 2\pi. \end{cases} \tag{4.69}$$

Here the diffraction coefficient is

$$D(\theta) = -\frac{\mathrm{e}^{\mathrm{i}\pi/4}}{2\sqrt{2\pi}}\left(\frac{1}{\cos(\frac{\theta+\alpha}{2})} + \frac{1}{\cos(\frac{\theta-\alpha}{2})}\right). \tag{4.70}$$

We notice that the field is split into three distinct regions, as shown in Figure 4.17. Region I contains the incident field $\mathrm{e}^{-\mathrm{i}kr\cos(\theta-\alpha)}$, reflected field $\mathrm{e}^{-\mathrm{i}kr\cos(\theta+\alpha)}$, and diffracted field $D(\theta)\frac{1}{\sqrt{kr}}\mathrm{e}^{\mathrm{i}kr}$. Region II contains the incident and diffracted fields. Region III contains the diffracted field alone. The diffracted field $D(\theta)\frac{1}{\sqrt{kr}}\mathrm{e}^{\mathrm{i}kr}$ has the form of a cylindrical wave

emanating from the edge tip with amplitude governed by the diffracted coefficient $D(\theta)$, as postulated by the GTD.

It is clear to see that the asymptotic approximation (4.69) is not uniform in θ, with a breakdown occuring in neigbourhoods of the shadow boundaries $\theta = \pi \pm \alpha$. In fact, this expansion is only valid when $\sqrt{kr}|\cos(\theta \pm \theta_0)| >> 1$. However we note that as $kr \to \infty$, these neighbourhoods become vanishingly small and the asymptotic approximation is valid over an ever increasing range of θ. Uniform asymptotic expressions for the diffraction coefficient $D(\theta)$ exist and hence this shortcoming of the GTD can be overcome. We direct the reader towards works on the Uniform Theory of Diffraction which was introduced in 1974 by Kouyoumjian and Pathak, see e.g. [86].

4.7.2   Diffraction by an impenetrable wedge

We now consider the diffraction of a time-harmonic plane wave $u^i = e^{-ikr\cos(\theta - \alpha)}$ by a wedge of exterior angle β with infinite planar faces. We wish to determine the field outside the wedge ($0 < \theta < \beta$) satisfying the Helmholtz equation and subject to the Dirichlet boundary condition $u = 0$ on Γ. The setup is shown in Figure 4.18. This problem was
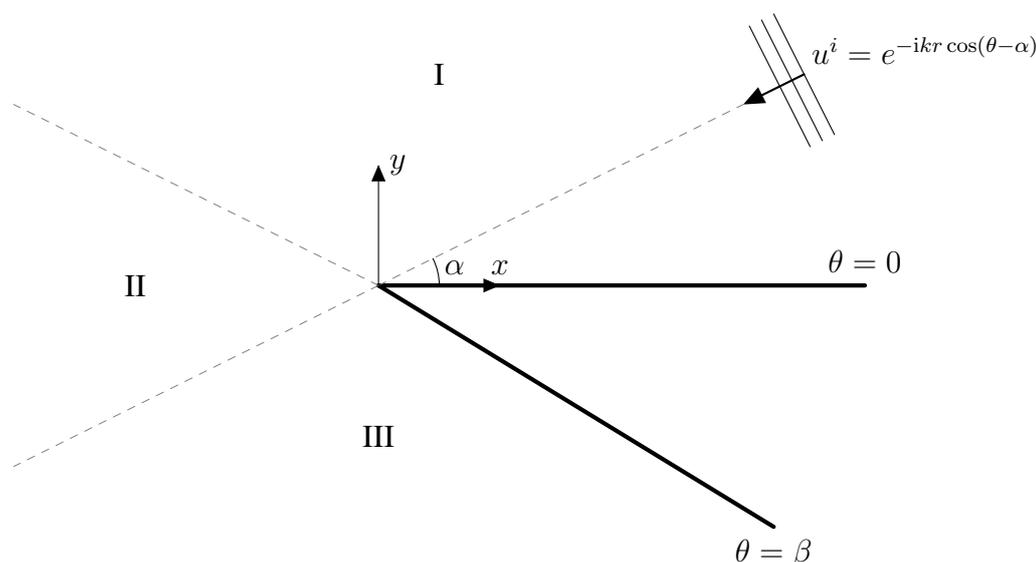
Figure 4.18: Scattering of a plane wave by a sound-soft infinite wedge.

first solved by Macdonald in 1902 [93] via the method of separation of variables. This was following the work of Sommerfeld in 1896 who first solved the problem of diffraction by a knife edge discussed in §4.7.1. Macdonald's solution was given as an infinite sum over Bessel functions which may then be converted to Sommerfeld integrals (see, e.g., [136])

convenient for asymptotic (large kr) analysis. We quote the final large kr result for the principal diffracted term from [136] as

$$u_d \sim D(\theta, \alpha, \beta) \frac{e^{\mathrm{i}(kr+\pi/4)}}{\sqrt{2\pi kr}}, \qquad (4.71)$$

where

$$D(\theta, \alpha, \beta) = \frac{\sin\frac{\pi}{n}}{n} \left(\frac{1}{\cos\frac{\pi}{n} - \cos\frac{\theta-\theta_0}{n}} - \frac{1}{\cos\frac{\pi}{n} - \cos\frac{\theta+\theta_0}{n}} \right) \qquad (4.72)$$

and $n = \beta/\pi$. Here we can identify $D(\theta, \alpha, \beta)$ as the diffraction coefficient and we see that the diffracted field has the phase $e^{\mathrm{i}kr}$ as postulated by the GTD. We note that in the case with Neumann boundary conditions, the diffraction coefficient is simply modified from (4.72) by replacing the subtraction of the two terms in parentheses by addition.

As we have seen in this section, there is a great deal of knowledge about the diffraction behaviour in the impenetrable case. In contrast, for the case of the penetrable wedge no such classical theory exists and diffraction coefficients are not known. There has, however, been some work into the asymptotic solution for the problem of diffraction by a penetrable wedge which we shall briefly discuss in the following section. Thereafter we shall consider the analogous problem in the time-domain in order to gain a better understanding of the structure of the diffracted field. Finally we shall use the findings from this to justify a heuristic extension of the GTD framework to the penetrable case. We shall also draw analogies to the problem of a point source above an interface between two different media.

4.8 Diffraction by a penetrable wedge

The canonical problem under consideration is now the diffraction of a time-harmonic plane wave propagating in a medium of wavenumber k_1 by an infinite wedge of a second medium of wavenumber $k_2 \neq k_1$, with the total field and its normal derivative being continuous across the interface between the two media. For simplicity of exposition, we restrict attention to the case $k_1, k_2 > 0$, and in particular to the case $0 < k_1 < k_2$ (although the case $0 < k_2 < k_1$ can be dealt with similarly).

This problem has proven to be intractable to date with no known exact closed-form or asymptotic solution. Nevertheless, many attempts have been made to obtain a solution, for a review of which the reader is referred to [90, 120]. Most of these have relied on using the Kontorovich-Lebedev technique [121], Maliuzhinets technique [24, 97], or the Wiener-Hopf technique [117]. These approaches have so far not led to exact closed-form solutions. Instead, many authors have sought approximate (i.e., not derived from the exact solution) asymptotic solutions valid for large kr, i.e., in the far-field. These approximations
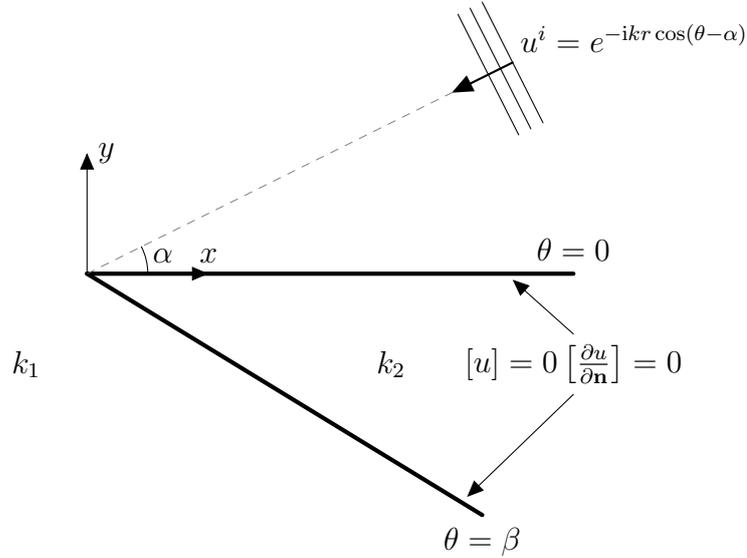
Figure 4.19: Scattering of a plane wave by an infinite penetrable wedge.

have been derived via approaches based on extensions of the GTD or Uniform Theory of Diffraction to the penetrable case [49, 77, 86, 121] and have been verified via comparisons to solutions obtained via numerical methods.

This extension to the penetrable case is to assume that the diffracted field emanating from the corner in Figure 4.19 comprises two partial cylindrical waves. The first radiates in the exterior region and takes the form

$$
u_d^{(1)} = \frac{\mathrm{e}^{\mathrm{i}k_1 r}}{\sqrt{k_1 r}} D_1(\theta, k_1) u^i,
$$

and the second radiates in the interior region and takes the form

$$
u_d^{(2)} = \frac{\mathrm{e}^{\mathrm{i}k_2 r}}{\sqrt{k_2 r}} D_2(\theta, k_2) u^i.
$$

This is the assumption we shall also make in this thesis. However, we expect non-uniformity in the diffraction coefficients D_1 and D_2 across shadow boundaries introduced by the GO. Further, we expect the presence of higher order terms (head waves), analogous to those observed in the problem of a point source above an interface between two different media (see §4.8.2). We now provide some justification of these statements by considering the associated time-domain problem.

4.8.1 Time-domain problem

As we have seen, little is known about the scattering of time-harmonic waves by a penetrable wedge, and indeed the same is true for time-dependent waves. However, if it is only

101

the phase structure of the field that we are interested in, we may appeal to the time-domain problem to glean some information, since here it is easier to use our intuition (and Huygen's principle) to better understand the behaviour of the scattered field. Moreover, there is a well-known correspondence (see, e.g., [20, pp. 349–355]) between the singularities of solutions of the time-dependent wave equation and the high frequency asymptotic behaviour of solutions of the Helmholtz equation. We summarise this correspondence as elucidated in [20] before analysing the scattering of a pulse by a penetrable wedge.

In the time domain, we seek a solution $U(\mathbf{x}, t)$ to the time-dependent wave equation

$$\frac{\partial^2 U(\mathbf{x}, t)}{\partial t^2} - c^2 \nabla^2 U(\mathbf{x}, t) = 0, \tag{4.73}$$

where c is the speed of sound in the fluid and t is time. $U(\mathbf{x}, t)$ is related to a solution of Helmholtz' equation via the Fourier transform with respect to time,

$$u(\mathbf{x}, k) := \int_{-\infty}^{\infty} U(\mathbf{x}, t) e^{ikt} dt. \tag{4.74}$$

The asymptotic behaviour of $u(\mathbf{x}, k)$ as $k \to \infty$ is governed by the singularities of $U(\mathbf{x}, t)$.

We suppose that a wavefield $U(\mathbf{x}, t)$ is the sum of waves $U_m(\mathbf{x}, t)$, each possessing a surface of discontinuity $t = \tau_m(\mathbf{x})$, which is a wavefront propagating with velocity c. For $t < \tau_m(\mathbf{x})$ we are ahead of the wavefront and we assume that $U_m = 0$, whereas close behind the wavefront, for $t > \tau_m(\mathbf{x})$ and $(t - \tau_m) \ll 1$, we assume that the wave may be expanded as the series

$$U_m(\mathbf{x}, t) \sim \sum_{n=0}^{\infty} (t - \tau_m(\mathbf{x}))^{\lambda_m + n} H(t - \tau_m(\mathbf{x})) A_n^{(m)}(\mathbf{x}), \tag{4.75}$$

where

$$H(t - \tau) = \begin{cases} 1, & t > \tau \\ 0, & t < \tau \end{cases}$$

is the Heaviside function and λ_m is a real constant which determines the smoothness in the vicinity of the wavefront. Then by formally taking a termwise Fourier transform of the expansion (4.75), we observe that each singularity $t = \tau_m(\mathbf{x})$ makes a contribution

$$\sum_{n=0}^{\infty} a_n \left(\frac{i}{k} \right)^{\lambda_m + n} A_n^{(m)}(\mathbf{x}) e^{ik\tau_m(\mathbf{x})}$$

to the asymptotic (large k) behaviour of $u(\mathbf{x}, k)$, for the appropriate constants a_n. This series is of the form of a ray expansion with τ_m satisfying the eikonal equation and $A_n^{(m)}$

satisfying the transport equations. This formal correspondence between singularities supported on wavefronts in the time-domain and high-frequency phase components in the frequency domain allows us to make qualitative predictions about the high-frequency wave behaviour for the penetrable wedge problem. To do so, we must consider the time-domain problem of diffraction of an incident plane pulse in a medium of wave speed c_1 by a wedge of wave speed c_2, with $0 < c_2 < c_1$. Here one can determine the position of the leading wavefronts associated with each of the components of the scattered field by appealing to Huygen's principle.

(a) $t < 0$ (b) $t > 0$
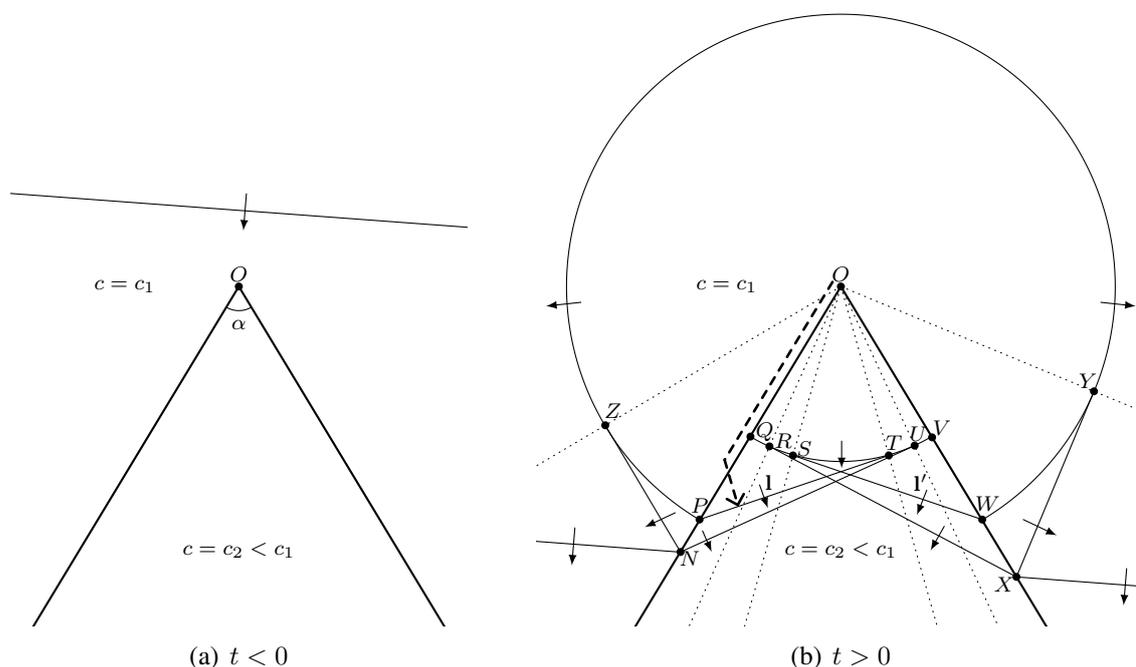
Figure 4.20: Wavefront diagrams for time-domain diffraction by a penetrable wedge, in the case where $c_2 < c_1$ and $\alpha > \cos^{-1}(c_2/c_1)$. The incident wavefront is assumed not to be in contact with the wedge for $t < 0$ and to arrive at the point O at time $t = 0$. The dotted lines in (b) indicate shadow boundaries, and the thick dashed arrow represents a ray path associated with the lateral wavefront PT.

An illustration of the resulting wavefront diagrams for one particular scattering configuration is shown in Figure 4.20. Here we have assumed that the incident wavefront is not in contact with the wedge before it reaches the diffracting corner (see Figure 4.20(a)). After it reaches the corner, the wavefront structure shown in Figure 4.20(b) emerges. The incident wavefront now has two components (intersecting the wedge at N and X in Figure 4.20(b)), and there exist two planar reflected wavefronts (NZ and XY) and two planar transmitted wavefronts (NU and XR). The diffracted wavefronts in the exterior and interior are segments of the circles centered at O of radius c_1t ($PZYW$) and c_2t ($QRSTUV$) respectively

103

(at time $t > 0$). In addition, Huygen's principle predicts the existence of so-called *head waves* (sometimes known as *lateral waves* or *bow waves*), with associated planar wavefronts (PT and WS). These waves can be associated with diffracted rays propagating along the exterior surface of the wedge at speed c_1, which shed new rays propagating into the interior medium. A typical ray path is shown as a thick dashed arrow in Figure 4.20(b). Similar waves also appear in the scattering of the field due to a point source by a planar interface (see, e.g., [22] and [76]). We note that Figure 4.20 shows only the simplest possible case, and more complicated wavefront configurations are possible. For example, for small enough wedge angles α the transmitted and head waves generated by one face of the wedge can be internally reflected by the other face, generating additional wavefronts. A sufficient and necessary condition for there to be no such internal reflection of the head waves is that $\alpha > \cos^{-1}(c_2/c_1)$.

In the original frequency domain wedge problem, we expect the structure of the field far from the corner to be analogous to that described above. That is, in the exterior we expect: an incident plane wave; two reflected plane wave beams, one bounded by the radial lines extending ON and OZ, and another bounded by the radial lines extending OX and OY; a diffracted wave with phase function $\mathrm{e}^{\mathrm{i}k_1 r}$, where r represents radial distance from the corner O. In the interior we expect: two transmitted plane wave beams; a diffracted wave with phase function $\mathrm{e}^{\mathrm{i}k_2 r}$; head waves with the phase functions $\mathrm{e}^{\mathrm{i}k_2 \mathbf{l}\cdot\mathbf{x}}$ and $\mathrm{e}^{\mathrm{i}k_2 \mathbf{l}'\cdot\mathbf{x}}$, where $\mathbf{l}, \mathbf{l}'$ are the direction vectors shown in Figure 4.20(b).

Although we have obtained some qualitative phase information of the wavefield through this approach, we do not know the nature of the singularities λ_m on each wavefront. Hence we do not know the relative magnitudes of the head waves to the diffracted circular waves. In order to gain some clues about the size of the head wave in relation to the diffracted waves, we look to a final related canonical problem: a point source at a plane interface between two media. This problem is of course quite different to the wedge problem, however it shares some important features, namely the head wave, and circular reflected and refracted waves which we may liken to the diffracted waves for the wedge.

4.8.2   Point source at an interface in the frequency domain

We consider the 3D problem of the scattering of a time-harmonic spherical wave from a planar boundary separating two different media with characteristic wavenumbers k_1, k_2. We shall take the planar boundary to be the plane $y = 0$ and also we shall only focus on the fields in the (x, y)-plane. Suppose the source of the spherical wave and the receiver are located in the upper medium ($y > 0$) with associated wavenumber k_2. The lower medium has the associated wavenumber k_1. We consider the specific case $k_1 < k_2$ (corresponding

to $c_1 > c_2$ in the time domain). This is the case for which the head wave will occur and be detected at the receiver. Across the interface, we impose the condition that the total field and its normal derivative are continuous. The reason for considering the 3D version is because that is the problem which has been covered in the literature (see, e.g., [38, §4]). The 2D problem should be amenable to the same treatment and should yield qualitatively similar results with the rates of decay suitably adjusted, although we shall not give details here.
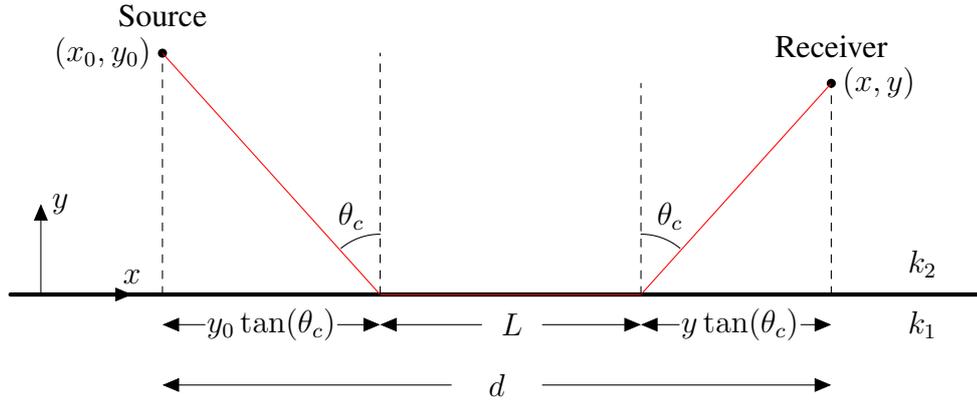
Figure 4.21: The propagation of a head wave due to total internal reflection.

Consider a point source

$$u^i(\mathbf{x}, k_2) = \frac{\mathrm{e}^{ik_2 r}}{4\pi r}; \quad r = \sqrt{(x - x_0)^2 + (y - y_0)^2},$$

located at the point $\mathbf{x}_0 = (x_0, y_0)$ in the upper medium and a receiver located at $\mathbf{x} = (x, y)$. This setup is displayed in Figure 4.21. Note that the source and receiver are located in the $z = 0$ plane, therefore we may neglect the z-direction in what is to follow. The total field received at $\mathbf{x} = (x, y)$ is

$$u(\mathbf{x}) \approx u^i(\mathbf{x}) + u^r(\mathbf{x}) + u^h(\mathbf{x})$$

where $u^r(\mathbf{x})$ is the expected reflected spherical wave (from the method of images) and $u^h(\mathbf{x})$ is the head wave. $u^r(\mathbf{x})$ is given simply as

$$u^r(\mathbf{x}) = A^r \frac{\mathrm{e}^{ik_2 R}}{4\pi R} \mathrm{e}^{ik_2 R},$$

where A^r is the reflection coefficient and $R = \sqrt{(x-x_0)^2 + (y+y_0)^2}$ is the distance from the image source to the receiver.

The head wave is found to have the asymptotic form [38, p.126]

$$u^h(\mathbf{x}) \sim \frac{\mathrm{i}}{2\pi} \frac{1}{(k_2/k_1)^2 - 1} \frac{\mathrm{e}^{\mathrm{i}[k_1 L + k_2(y_0 + y)\sec\theta_c]}}{k_1 d^{1/2} L^{3/2}}, \quad \text{as } k_2 d \to \infty$$

where $d = x - x_0$, $L = d - (y_0 + y)\tan\theta_c$, and $\theta_c^{-1} = \sin^{-1}(k_1/k_2)$ is the critical angle related to total internal reflection. Thus, assuming k_2/k_1 is fixed,

$$\left| \frac{u^h}{u^r} \right| = \mathcal{O}\left(\frac{1}{k_2 d} \right), \quad \text{as } k_2 d \to \infty,$$

assuming d/L and R/r are both $\mathcal{O}(1)$. Hence, the magnitude of the head wave is $\mathcal{O}(1/(k_2 d))$ smaller than that of the incident wave.

In order to gain a better physical understanding of the origin of the head wave, it is instructive to pictorially consider the time domain version of the problem. This is shown in Figure 4.22 where we consider the slice through the (x, y)-plane containing the source. We see that the wavenumbers k_1 and k_2 have been replaced by their corresponding wavespeeds c_1 and c_2 where here $c_1 > c_2$, i.e., the waves propagate faster in the lower medium with speed c_1. The head wave can be seen as being composed of those rays that strike the inter-
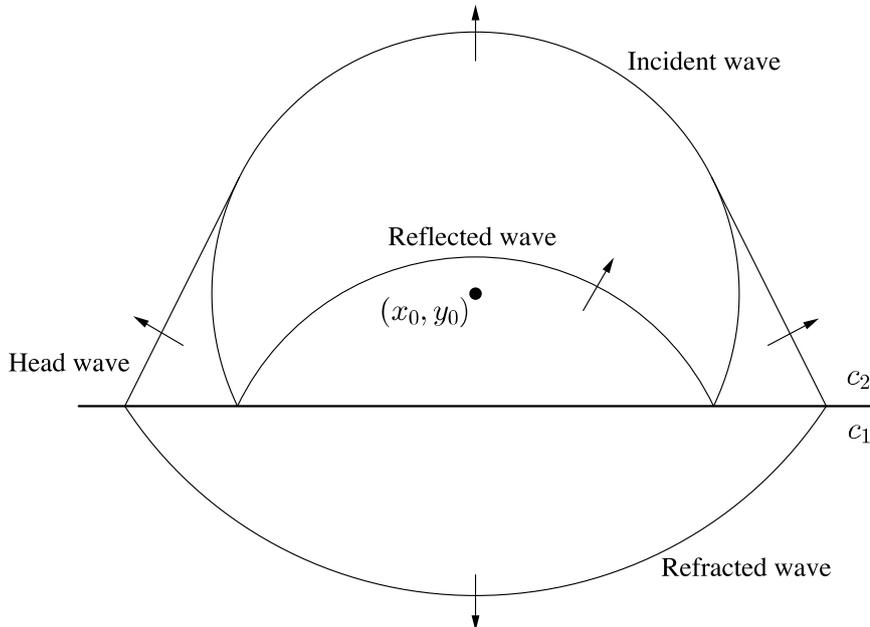
Figure 4.22: The propagation of a head wave due to total internal reflection.

face at an angle greater than the critical angle θ_c (corresponding to total internal reflection)

106

propagate along the boundary for a distance and are then shed at the angle θ_c back into the more optically dense medium. Hence the head wave is a plane wave travelling in direction $(\sin\theta_c, \cos\theta_c)$ with a phase lag corresponding to the distance L. In the time domain problem, the head wave connects the two wavefronts corresponding to the incident and refracted waves.

Chapter 5

Hybrid Numerical-Asymptotic approximation

In this chapter we develop our HNA approximation space for the problem of high-frequency scattering by a penetrable convex polygon. The starting point will be to decompose the unknown $v = (u, \partial u / \partial \mathbf{n})$ in our integral equation formulation as

$$v(\mathbf{x}) = v_{go}(\mathbf{x}) + v_d(\mathbf{x}), \quad \mathbf{x} \in \Gamma,$$

where $v_{go} = (u_{go}, \partial u_{go} / \partial \mathbf{n})$ is the GO approximation to v, and the remainder $v_d = (u_d, \partial u_d / \partial \mathbf{n})$ is interpreted as the diffracted field. The GO approximation v_{go} represents the leading-order behaviour at high frequencies and its calculation is performed via the beam tracing algorithm detailed in the previous chapter. We aim to approximate the diffracted field using an ansatz of the form

$$v_d(\mathbf{x}) \approx \sum_{m=1}^{M_1} v_{1,m}(\mathbf{x}, k_1) \, \exp(\mathrm{i} k_1 \psi_{1,m}(\mathbf{x})) + \sum_{m=1}^{M_2} v_{2,m}(\mathbf{x}, k_2) \, \exp(\mathrm{i} k_2 \psi_{2,m}(\mathbf{x})), \quad (5.1)$$

which generalises the standard HNA ansatz (1.1) to the case where two different wavenumbers are present. This chapter focuses on the choice of suitable phases $\psi_{j,m}$ based on the qualitative analysis of the penetrable wedge diffraction problem presented in §4.8. In principle, in order to completely capture the oscillatory behaviour we would have to take into account

(i) the diffracted and head waves emanating from each corner of the polygon,

(ii) the (infinitely many) internal re-reflections of these waves.

In order to achieve a highly accurate solution we would need to include a large number of these phases in the ansatz (5.1). However, this would lead to a potentially high computational cost and a complicated implementation. Instead, our approach will be to include

only a small number of the most important phases so as to obtain a computationally cheap method which may be implemented relatively simply. More precisely,

> we aim to achieve "engineering accuracy" (as defined in §1.4) or better with a small number of degrees of freedom fixed independent of the frequency.

The phrase "engineering accuracy" was discussed in §1.4, and is defined here as approximately 1% relative error in the solution in either scattering domain Ω_1 or Ω_2 or in the far-field. To obtain such an accuracy in these regions, experience suggests that an accuracy of 2% relative error or better on the boundary should be sufficient.

We emphasise that this is a *chosen* error tolerance which leads to the specific HNA approximation space developed in this thesis. In principle, as we discuss in §5.4 and §5.5, a higher accuracy of approximation could be achieved by modifying our approximation space to include further phase functions, at the expense of increased computational cost and implementation complexity. However, certain tolerances (to be described later) which dictate aspects of the meshing and inclusion of certain basis functions would have to be adjusted accordingly. These tolerances will be highlighted as we come to them.

In order to construct our HNA approximation, we begin by proposing in §5.2 two simple ansatzes of the form (5.1) which aim to capture the leading order behaviour of v_d. Both ansatzes incorporate phases relating to the diffracted waves emanating from each of the corners of the polygon. Phases associated with the head waves and multiple re-reflections of diffracted and head waves are not included.

The efficacy of approximation spaces based on these ansatzes will be tested via a best fitting procedure to be described in §5.3.1. The results obtained through this process will inform the construction of a final ansatz which is a combination of the initial two proposed. A Galerkin BEM implementation based on an approximation space derived from this final ansatz will be described in §6.

To provide some context for what follows, we begin by reviewing the approach taken in [30, 63] for the analogous impenetrable problem of scattering by a sound-soft convex polygon.

5.1 The sound-soft case

The sound-soft scattering problem is characterised by the Dirichlet boundary condition $u = 0$ on the scatterer's boundary Γ. Here the unknown boundary data is $\partial u/\partial \mathbf{n}$ alone since u is specified by the boundary condition. When the scatterer is a convex polygon, the

HNA ansatz (1.1) (which involves only one wavenumber, call it k) contains just two terms in the summation, so that we have on a typical side Γ_j of length L_j

$$\frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}(s)) = \frac{\partial u_{go}}{\partial \mathbf{n}}(\mathbf{x}(s)) + v^+(s)\mathrm{e}^{\mathrm{i}ks} + v^-(L_j - s)\mathrm{e}^{-\mathrm{i}ks}, \quad s \in [0, L_j], \tag{5.2}$$

i.e., with phases $\psi^\pm(\mathbf{x}(s)) = \pm s$, where s is arc length measured anti-clockwise around Γ. The geometrical optics term here is simply given by

$$\frac{\partial u_{go}}{\partial \mathbf{n}}(\mathbf{x}(s)) = \begin{cases} 2\frac{\partial u^i}{\partial \mathbf{n}}, & \text{if } \mathbf{d}^i \cdot \mathbf{n} < 0 \text{ (i.e., if } \Gamma_j \text{ is illuminated by the incident wave),} \\ 0, & \text{otherwise (i.e., if } \Gamma_j \text{ is in shadow).} \end{cases} \tag{5.3}$$

The phases $\psi^+ = s$ and $\psi^-(s) = -s$ correspond respectively to diffracted waves travelling anticlockwise and clockwise around the boundary. It is proven rigorously in [63] that the amplitudes $v^\pm$ are analytic in the right half-plane $\mathrm{Re}(s) > 0$, where they satisfy the following bounds, which we quote from [61] where they are expressed in a simpler (albeit slightly weaker) form than in [63]:

$$|v^\pm(s)| \leq \begin{cases} Ck^2|ks|^{-\delta_j^\pm}, & 0 < |s| \leq 1/k, \\ Ck^2|ks|^{-1/2}, & |s| > 1/k. \end{cases} \tag{5.4}$$

Here $\delta_j^\pm \in (0, 1/2)$ and C depend only the corner angles. As explained in [62, Remark 4.2], the bounds (5.4) imply that $v^\pm(s)$ are non-oscillatory since, by the Cauchy integral formula, it may be shown that the derivatives of these functions grow no faster than the functions themselves with respect to increasing k. Therefore, the functions $v^\pm(s)$ may be efficiently approximated by low-order piecewise polynomials on meshes appropriately graded towards the singularities at the corners. The main result of [63] demonstrates that, in order to maintain a prescribed accuracy as k increases, the number of degrees of freedom in the approximation space must increase at most at a rate of $\log^2(k)$. Accompanying numerical results suggest that, in fact, the number of degrees of freedom may be kept fixed. Similar results have been obtained for scattering by screens and apertures [62] and a class of non-convex polygons [29].

The remarkable success of the HNA methodology in the impenetrable case is due to two factors. Firstly, the high frequency asymptotic behaviour of the solution to the canonical problem of diffraction by an infinite impenetrable wedge is known - in fact there is an exact closed-form solution available which was reviewed in §4.7.2. This allows one to pick out the phases required to capture the primary diffracted waves. Secondly, the only multiple scattering effects in this case are the multiply-diffracted waves propagating around the boundary of the polygon. But each of these waves has one of the same two phases $\psi^\pm$

already included in the approximation space, so their contribution can be captured in the amplitudes $v^{\pm}$.

By contrast (as discussed in §4), no exact (or even asymptotic) solution has yet been derived for the analogous canonical problem of diffraction by a penetrable wedge. Furthermore, for the penetrable case the multiply-scattered field is extremely complicated, featuring multiple reflections/refractions of the incident and diffracted fields, with potentially infinitely many different phases to consider in the approximation of v_d. But using the intuition gained from the study of the time domain wedge problem in §4.8, we now show how we can develop some simple generalisations of the ansatz (1.1) to the penetrable case.

5.2 Hybrid numerical-asymptotic approximation space

In this section we describe perhaps the two simplest generalisations of the ansatz (1.1) to the penetrable case. The first contains the same two phases as in (1.1) but now with the addition of the interior wavenumber's counterparts so that in total, we have four phase functions. The second generalisation adds to this by also including waves diffracted from each corner that traverse the interior of the polygon and impinge on sides non-adjacent to that corner. We shall term the two approximation spaces generated from these ansatzes as "Approximation Space 1" (AS1) and "Approximation Space 2" (AS2).

The same two approximation spaces were presented by Groth et al. in [54] except for some small differences concerning the graded meshes and polynomial degrees employed. The methodology presented here stems from that paper however represents a significant improvement as the results later in the chapter shall demonstrate. We shall highlight the differences to the work in [54] as they become relevant and we take care to emphasise the important extensions to that work which have been incorporated in the present algorithm.

5.2.1 Approximation Space 1: including diffraction from adjacent corners

On each side of the polygon we include phases corresponding to diffracted waves emanating from the corners adjacent to that side. That is, in the ansatz (5.1) consists of only the four phases

$$\psi_{1,1} = s, \quad \psi_{1,2} = -s,$$
$$\psi_{2,1} = s, \quad \psi_{2,2} = -s.$$

More explicitly, this ansatz has the form on a typical side Γ_j of length L_j

$$v_d(\mathbf{x}(s)) \approx v_{1,1}(s)e^{ik_1s} + v_{1,2}(s)e^{-ik_1s} + v_{2,1}(s)e^{ik_2s} + v_{2,2}(s)e^{-ik_2s}, \quad \mathbf{x} \in \Gamma_j. \quad (5.5)$$

Here $v_{1,1}$, $v_{1,2}$, $v_{2,1}$, $v_{2,2}$ are amplitude functions which will be approximated by piecewise polynomials supported on overlapping graded meshes, designed to capture the expected singular behaviour at the corners of the polygon. More precisely, the meshes for $v_{1,1}$ and $v_{2,1}$ are graded towards $s = 0$, and the meshes for $v_{1,2}$ and $v_{2,2}$ are graded towards $s = L_j$ as depicted in Figure 5.1. Hence the approximation space is the set of functions

$$\{P_i(s)e^{ik_1s}, P_i(s)e^{-ik_1s}, P_i(s)e^{ik_2s}, P_i(s)e^{-ik_2s}\}, \quad \text{for } i = 0, \ldots, n,$$

where P_i are the aforementioned piecewise polynomials and n is the highest degree we wish to include.

To describe the graded meshes we employ, we consider the case of a geometric mesh on the interval $[0, L]$, $L > 0$, refined towards 0. The meshes for approximating $v_{1,1}$, $v_{1,2}$, $v_{2,1}$, $v_{2,2}$ on each side of the polygon are constructed from this basic building block by straightforward coordinate transformations. Given $n \geq 1$ (the number of layers in the mesh) we let $G_n(0, L)$ denote the set of meshpoints $\{x_i\}_{i=0}^n$ defined by

$$x_0 := 0, \quad x_i := \sigma^{n-i}L, \quad i = 1, 2, \ldots, n,$$

where $0 < \sigma < 1$ is a grading parameter. A smaller grading parameter represents a more severe grading. Here, based on numerical experimentation, we choose a grading parameter of $\sigma_1 = 0.2$ to approximate $v_{1,1}$, $v_{1,2}$ and a grading parameter of $\sigma_2 = 0.19$ to approximate $v_{2,1}$, $v_{2,2}$. We note that many other choices give good approximations, for example, in [54] by Groth et al., the choice $\sigma_1 = \sigma_2 = 0.15$ is made yielding good results. It was found, however, in performing the boundary element implementation of the proposed approximation spaces (the topic of §6), that shifting the graded meshes by choosing $\sigma_1 \neq \sigma_2$ was beneficial in terms of conditioning. The precise values $\sigma_1 = 0.2$, $\sigma_2 = 0.19$ were chosen after some numerical experiments but yielded only marginally better results than other choices such as $\sigma_1 = 0.17$, $\sigma_2 = 0.15$, for example. We shall postpone briefly the discussion of the reason for the occurrence of poor conditioning for $\sigma_1 = \sigma_2$ until the end of this subsection and first define the meshes we shall use.

Given a vector $\mathbf{p} \in (\mathbb{N}_0)^n$, we let $P_{\mathbf{p},n}(0, L)$ denote the space of piecewise polynomials on the mesh $G_n(0, L)$ with the degree vector $\mathbf{p}$, i.e.,

$$P_{\mathbf{p},n}(0, L) := \{\rho : [0, L] \to \mathbb{C} :$$

$$\rho|_{(x_{i-1}, x_i)} \text{ is a polynomial of degree less than or equal to } (\mathbf{p})_i, \ i = 1, \ldots, n\}.$$

For reasons of efficiency and conditioning it is common to decrease the order of the approximating polynomials towards the singularity. Specifically, in all of our experiments we use a "linear slope" degree vector $\mathbf{p}$ with

$$(\mathbf{p})_i := \begin{cases} p - \left\lfloor \frac{(n+1-i)}{n}p \right\rfloor, & 1 \leq i \leq n-1, \\ p, & i = n, \end{cases}$$

where the integer $p \geq 0$ is the highest polynomial degree on the mesh. The $\lfloor \cdot \rfloor$ symbol denotes the floor function which is defined for a real number x as

$$\lfloor x \rfloor = \max\{m \in \mathbb{Z} : m \leq x\}.$$

For simplicity we assume the same number of layers, n, in each of the graded meshes on the polygon. We adopt an "hp" refinement approach (as in [63]) in which the number of degrees of freedom is increased by increasing the polynomial degree p, while simultaneously refining the meshes. Specifically, in all our experiments we take

$$n = \lceil C(p+1) \rceil \tag{5.6}$$

with $C = 3/2$. The $\lceil \cdot \rceil$ symbol denotes the ceiling function which is defined for a real number x as

$$\lceil x \rceil = \min\{m \in \mathbb{Z} : m \geq x\}.$$

In [54] this constant was taken to be $C = 1$, however it was later found by the author that $C = 3/2$ yields marginally better results.
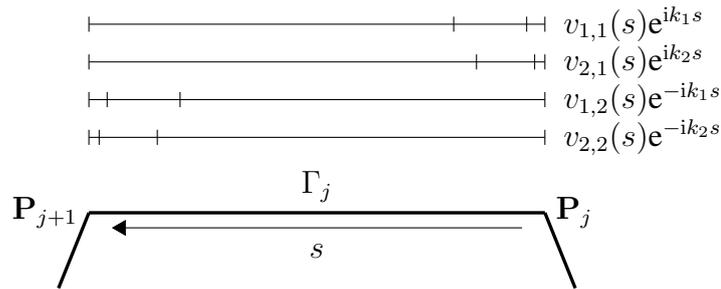
Figure 5.1: Illustration of overlapping geometrically graded meshes used to approximate the amplitudes $v_{1,1}$, $v_{1,2}$, $v_{2,1}$, $v_{2,2}$ associated with the phase functions on a typical side Γ_j.

Conditioning problems for the choice $\sigma_1 = \sigma_2$

It was mentioned that the above described system of overlapping meshes may lead to ill-conditioning in the boundary element implementation when choosing $\sigma_1 = \sigma_2$, i.e., when exactly the same mesh is used for the k_1 and k_2 basis functions. Ill-conditioning

arises when the wavelengths $\lambda_1 = 2\pi/k_1$ and $\lambda_2 = 2\pi/k_2$ are of a comparable length to some of the elements. On these small elements, the polynomials multiplying the oscillatory functions $e^{ik_1 s}$ and $e^{ik_2 s}$ (or $e^{-ik_1 s}$ and $e^{-ik_2 s}$) may be of a high enough degree to resolve the difference in oscillation between these two functions. Hence, some of the basis functions become redundant and the linear system ill-conditioned.

A systematic way to avoid redundancy would be to stop increasing p for one of $e^{ik_1 s}$ and $e^{ik_2 s}$ once ill-conditioning starts to arise. Specifically, on an element of length h, one might stop increasing p for $e^{ik_2 s}$ (for example) once the quantity

$$\frac{4\pi(p+1)}{h(k_2 - k_1)},$$

which represents the #DOF on the element divided by the wavelength of the wave $e^{i(k_2 - k_1)s}$, exceeds a certain tolerance. Of course if k_2 is complex, we replace k_2 in the above with $\mathrm{Re} k_2$. However, this is *not* the approach implemented in this thesis. Instead, it was found that adopting the less systematic, but simpler, remedy of taking $\sigma_1 \neq \sigma_2$ was sufficient to avoid the conditioning problems described above, and hence this latter approach is the one implemented here.

5.2.2   Approximation Space 2:
including diffraction from non-adjacent corners

The second ansatz we propose builds on the first by including extra phases corresponding to the diffracted waves which traverse the interior of the scatterer up to their first reflection. This equates to an ansatz with phases

$$\psi_{1,1} = s, \quad \psi_{1,2} = -s,$$
$$\psi_{2,1} = s, \quad \psi_{2,2} = -s,$$
$$\psi_{2,3} = r_1, \quad \psi_{2,4} = r_2, \ldots, \psi_{2,n_s} = r_{n_s - 2},$$

where $r_i, i = 1, \ldots n_s - 2$ are the radial distances from the corners non-adjacent to the side in question (there are $n_s - 2$ such corners since the polygon is convex). Written explicitly, the ansatz on the side in question has the form

$$v_d(\mathbf{x}) \approx v_{1,1}(s(\mathbf{x}))e^{ik_1 s(\mathbf{x})} + v_{1,2}(s(\mathbf{x}))e^{-ik_1 s(\mathbf{x})} + v_{2,1}(s(\mathbf{x}))e^{ik_2 s(\mathbf{x})} + v_{2,2}(s(\mathbf{x}))e^{-ik_2 s(\mathbf{x})} +$$
$$+ v_{2,3}(r_1(\mathbf{x}))e^{ik_2 r_1(\mathbf{x})} + v_{2,4}(r_2(\mathbf{x}))e^{ik_2 r_2(\mathbf{x})} + \ldots + v_{2,n_s}(r_{n_s - 2}(\mathbf{x}))e^{ik_2 r_{n_s - 2}(\mathbf{x})},$$
$$(5.7)$$

where the additional amplitude functions $v_{2,i+2}, i = 1, \ldots, n_s - 2$ are approximated numerically by piecewise polynomials.
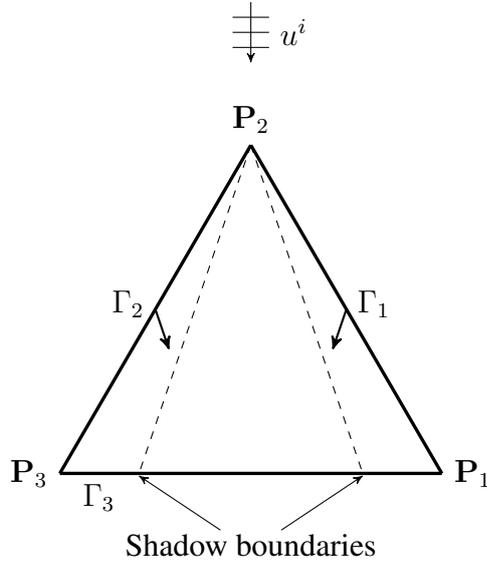
Figure 5.2: Shadow boundaries in an equilateral triangle for $\mathbf{d}^i = (0, -1)$. Determining the mesh on Γ_3 associated with the approximation of $v_{2,3}$. Mesh points are introduced on Γ_3 at the locations of the shadow boundaries associated with the primary transmitted waves from sides Γ_1 and Γ_2. The resulting mesh has three elements

We expect the amplitudes $v_{2,i+2}$ to have a (possibly large) number of discontinuities to compensate for the discontinuities inherent in the GO approximation (where we cut off the plane wave beams sharply across the beam boundaries, see §4.2). In principle, one should therefore approximate each $v_{2,i+2}$ on a mesh refined towards each of these discontinuities, as discussed in [61]. However, for simplicity we shall (to begin with) take into account only those discontinuities arising from the lowest order GO terms, i.e., primary transmitted waves (later, in §5.3.3, we will discuss the treatment of shadow boundaries associated with higher order GO terms). At these discontinuities we introduce a single mesh point rather than introducing a multiple element refined mesh, again for simplicity. Let us consider an example to make clear precisely which shadow boundaries we take into account in AS2. Consider the scattering of a plane wave with $\mathbf{d}^i = (0, -1)$ by an equilateral triangle as shown in Figure 3.4. Figure 5.2 shows the locations of the shadow boundaries on Γ_3 arising from the primary transmitted beams from sides Γ_1 and Γ_2. For the approximation of $v_{2,3}$ on this side we place a mesh point at each of these locations here producing a mesh with three elements. The resulting meshes on side Γ_3 for this problem are shown in Figure 5.3.

The triangle example is the simplest since, for each side, we have only a single basis function of the form $e^{ik_2 r}$. For a general, n_s-sided polygon, we will have $n_s - 2$ such basis functions on each side. In this more general case, we shall accommodate all of these basis functions on a single mesh which is constructed in the following way. We begin with a
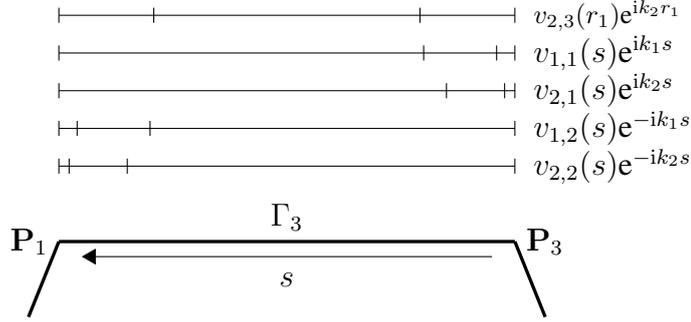
Figure 5.3: Illustration of overlapping meshes used to approximate the amplitudes $v_{1,1}$, $v_{2,1}$, $v_{1,2}$, $v_{2,2}$, $v_{2,1}$ and $v_{2,3}$ associated with the phase functions on side Γ_3 for the triangle example considered in Figure 5.2.
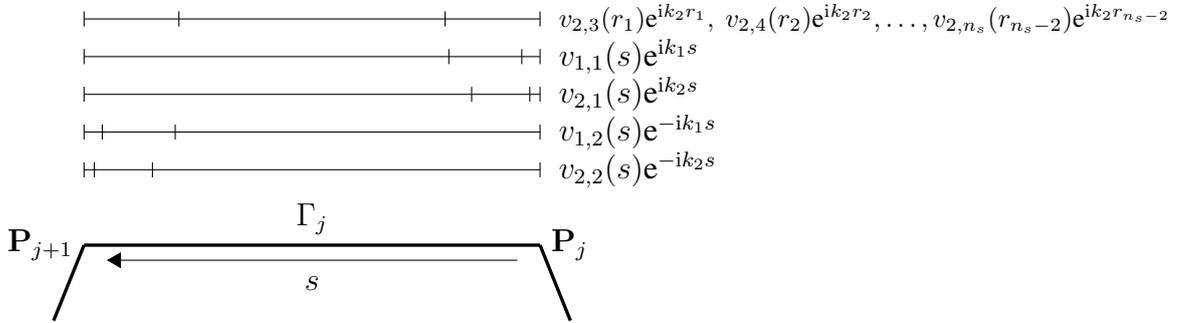
Figure 5.4: Illustration of overlapping geometrically graded meshes used to approximate the amplitudes $v_{1,1}$, $v_{2,1}$, $v_{1,2}$, $v_{2,2}$, $v_{2,1}$ and $v_{2,3}, \ldots, v_{2,n_s}$ associated with the phase functions on the side Γ_j for an n_s-sided convex polygon.

single element spanning the whole side Γ_j. Then if (during the beam tracing algorithm for computing the GO term) any shadow boundaries from primary transmitted beams intersect the side Γ_j, we insert new mesh points at the locations of these intersections. The resulting sequence of meshes for a polygon with n_s sides is shown in Figure 5.4. On each of these resulting elements we approximate each of $v_{2,i}$, $i = 3, \ldots, n_s$, by a single polynomial of degree p, where p is the same as for the AS1 amplitudes. This process adds at most $(2n_s - 3)(p + 1)$ degrees of freedom on the side Γ_j; extending this procedure to all of the other sides results in at most $n_s(2n_s - 3)(p + 1)$ degrees of freedom begin added when we go from AS1 to AS2.

We remark that it may be more natural to approximate (on a side Γ_j) each of the amplitudes on its own mesh leading to a large sequence of overlapping meshes. However, here the single mesh approach was chosen for its relative ease of implementation, and a study of this method with overlapping meshes is left to future work.

5.2.3 Including other phase functions

We expect that even more accurate approximations could be obtained by including the effects of higher order terms in the asymptotic approximation. Firstly, one could include phases associated with the head waves associated with each corner of the polygon. Secondly, one could include phases associated with the (multiple) internal reflection of (i) the diffracted waves and (ii) the head waves. The phases for (i) could be computed using an image method (i.e. introducing "image corners" in a non-physical image domain outside the scatterer). The phases for (ii) could be determined using a simple modification of the beam-tracing algorithm described in §4.2. However, we do not consider these generalisations any further here since, as we shall demonstrate shortly, for the range of problems we consider, the approximation spaces 1 and 2 are sufficient to achieve our goal of engineering accuracy with a fixed number of DOF.

5.3 Testing the HNA approximation spaces 1 and 2

In the previous section, two approximation spaces ("Approximation Space 1", defined in §5.2.1, and "Approximation Space 2", defined in §5.2.2) were proposed for the approximation of v_d. In this section, we assess, via numerous numerical examples, the efficacy of AS1 and AS2 to accurately approximate v using just a small number of degrees of freedom. This is done by best fitting each approximation space to a reference solution obtained using the conventional hp-BEM described in Chapter 3 and hence obtaining the numerical "best approximation" possible from each approximation space. This will validate the phase choices made in the previous section and identify for which problem parameters (absorption, incident wave direction etc.) the approximation spaces are effective. For each example considered below, we first obtain the reference solution $v_{\text{ref}} \approx v$ by solving our system of BIEs (2.98) using the conventional hp-BEM with a sufficient number of degrees of freedom to ensure that the relative error

$$\|v - v_{\text{ref}}\|_{L^2(\Gamma)} / \|v\|_{L^2(\Gamma)}$$

is 5×10^{-5} or less. Henceforth, for ease of presentation we shall denote this reference solution v_{ref} simply as v. Next, we compute an approximation to $v_{go} = (u_{go}, \partial u_{go}/\partial \mathbf{n})$ using the BTA described in §4 with our "rules of thumb" (4.55) and (4.61). Finally, a least squares approach is employed to find the best fit from each of AS1 and AS2 to $v_d = v - v_{go}$ in the $L^2(\Gamma)$ norm. This is carried out by discretising the $L^2(\Gamma)$ norm to be minimised using a large number of equally spaced quadrature points on each side, and solving the resulting

discrete least squares problem as described in §5.3.1 below. We denote the approximation to $v = (u, \partial u/\partial \mathbf{n})$ achieved via this procedure using Approximation Space j by $V_j = (U_j, W_j)$, $j = 1, 2$.

The least squares procedure is also useful for tuning certain parameters which define the approximation space, such as the mesh grading parameters (σ_1 and σ_2), maximum polynomial degree (p) and polynomial degree variation across the meshes (C). Much of the tuning in these parameters we shall omit to describe for sake of brevity. However, it should be noted that there are many different choices available which lead to good approximations. The particular choices made here result from extensive numerical experiments and the experience of the author. The parameter choice which we shall discuss in detail is that of the maximum polynomial degree p. The maximum polynomial degree p is taken as 3 initially (based on preliminary experiments) but we shall examine the convergence of the method in p in §5.7 where it becomes evident that high p (> 3) provides little improvement over $p = 2$ or $p = 3$ for many problems (except for those of high absorption).

5.3.1   Best approximation via least squares

Least squares data fitting is a well-known tool, nevertheless we describe it here briefly before presenting results based on the two approximation spaces. We commence by stating the least squares problem in general terms before stating it more specifically for the problem at hand. The (linear) least squares problem is the following [134]:

$$\begin{aligned} &\text{Given } \mathbf{A} \in \mathbb{C}^{m \times n}, m \geq n, \mathbf{b} \in \mathbb{C}^m, \\ &\text{find } \mathbf{x} \in \mathbf{C}^n \text{ such that } ||\mathbf{b} - \mathbf{A}\mathbf{x}||_2 \text{ is minimised,} \end{aligned} \qquad (5.8)$$

where $|| \cdot ||_2$ denotes the 2-norm, that is, the discrete version of the L^2-norm. The 2-norm corresponds to the Euclidean distance, so we can interpret this statement geometrically. We seek a vector $\mathbf{x} \in \mathbb{C}^n$ such that $\mathbf{A}\mathbf{x} \in \mathbb{C}^m$ is the closest point in range($\mathbf{A}$) to $\mathbf{b}$.

In the specific problem of interest here, each of the n columns of $\mathbf{A}$ corresponds to one of the n basis functions in our approximation space. The m rows correspond to m points $\{s_i\}_{i=1}^n$ around the boundary Γ at which each basis function is evaluated. The points $\{s_i\}_{i=1}^n$ should be chosen such that the 2-norm approximates the L^2-norm on Γ accurately. It was found that choosing s_i evenly spaced with $m = 60\mathrm{Re}(\mathrm{k}_2)\mathrm{n}_\mathrm{s}$, i.e., 60 evaluation points per wavelength $\lambda_2 = 2\pi/\mathrm{Re}(\mathrm{k}_2)$, is sufficient to achieve close to machine precision in this approximation. We note that $m > n$ always and hence the system $\mathbf{A}\mathbf{x} = \mathbf{b}$ is overdetermined. The vector $\mathbf{b}$ is v_d evaluated at $\{s_i\}_{i=1}^n$.

Our least squares problem (5.8) may now be solved to find $\mathbf{x} \in \mathbf{C}^n$. There are numerous possible approaches for this task. Here we solve the least squares problem by simply typing `x=A\b` in Matlab which employs QR factorisation.

5.3.2   Numerical examples

We shall test and develop the approximation spaces 1 and 2 mainly by considering the example of scattering by an equilateral triangle with side length 2π. This is done for two main reasons. Firstly, the triangle is the polygon with the fewest corners and so offers the best chance of isolating the leading order diffraction terms we are trying to study. Secondly, the 3D counterpart, the trigonal polyhedron, is of interest to the application of scattering by atmospheric ice crystals as discussed in Chapter 1 and [110]. By testing on the simple triangle case, we aim to determine the regimes of k_1 and n_I for which each approximation space is valid, and also to alter any aspects of the spaces which are leading to poor approximation. It shall be seen that the main cause of error is the presence of the discontinuities introduced by subtracting off v_{go} which is discontinuous across shadow boundaries. Using experiments, we determine parameter-based rules which introduce additional mesh points at these shadow boundaries provided certain tolerances are exceeded. Finally, we develop a finalised approximation space, called Approximation Space 3, which incorporates Approximation Space 1 and Approximation Space 2, and the additional shadow boundary rules.

We shall consider two different incident angles, as shown in Figure 3.4 and Figure 5.5, and we consider six different levels of absorption, governed by the imaginary part of the refractive index $n = n_R + i n_I$. Specifically, the real part of the refractive index is taken to be 1.5 throughout, which, as discussed in Chapter 1, represents a rough average of the varying (with wavelength) refractive index of ice. The imaginary part takes the values $n_I = 1/20, 1/40, 1/80, 1/160, 1/320, 0$. Throughout we take $\alpha = 1$.

It ought to be noted that the work in this section is similar to that of the present author (and colleagues) in [54] in which the same two approximation spaces are considered and tested via a best-fitting approach. There are some key differences, however, between the work here and that presented in [54]. In [54] numerical experiments are presented concerning scattering by an equilateral triangle, as here, but with a refractive index with $n_I = 1/20, 1/40, 1/80, 0$, i.e., at four different levels of absorption, and with $n_R = 1.31$. The value $n_R = 1.31$ was chosen since it corresponds to the real part of the refractive index of ice at a wavelength of $0.55\mu m$. Four different incident wave angles are considered: the two considered here in Figure 3.4 plus two evenly spaced intermediate angles. The beam tracing algorithm used in [54] is identical to that used in this thesis except for the crucial fact that it employs the choice GO1 throughout (see §4.3.2 and §4.3.4 of this thesis)
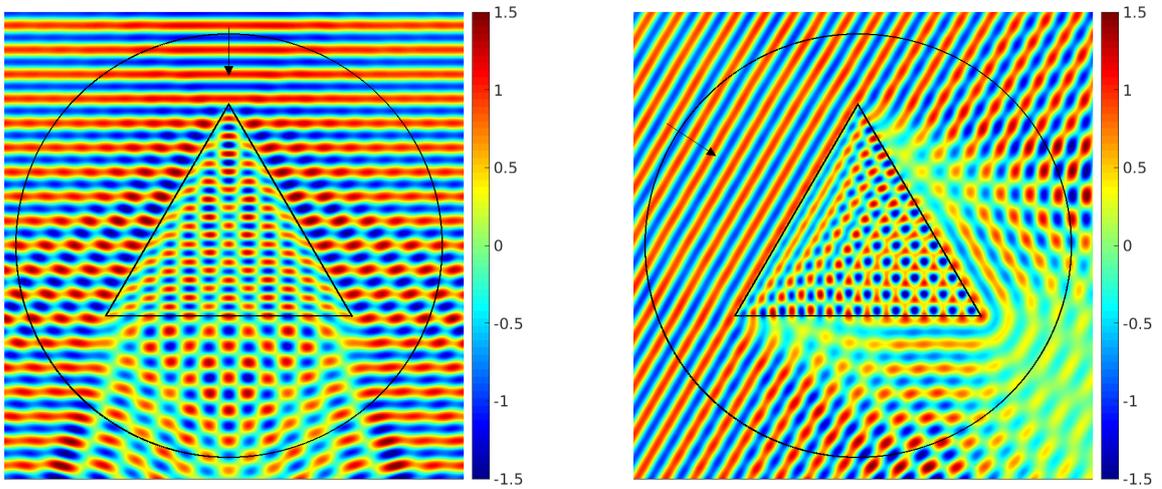
(a) $\mathbf{d}^i = (0, -1)$.   (b) $\mathbf{d}^i = (\cos(\pi/6), -\sin(\pi/6))$.

Figure 5.5: Scattering of a plane wave with $k_1 = 10$ by the triangle in Figure 3.4 with $n = 1.5 + 0.00625i$.

which leads to an extremely poor GO approximation at certain incident angles, the worst GO approximation occuring at the angle corresponding to Angle 2 here. This leads to poor performance of the HNA approximation spaces at this angle. For the incident angle corresponding to Angle 1 here, however, the approximation spaces perform well. For this angle, AS1 (with 360 degrees of freedom) yields accuracies of better than 3% on the boundary provided $k_1 n_I \geq 1$, and approximation space 2 (with 410 degrees of freedom) yields accuracies of approximately 3% or better provided $k_1 n_I \geq 0.5$. The grading parameters chosen in [54] for the graded meshes were taken as $\sigma_1 = \sigma_2$. C was taken as 1 and p fixed at 4.

Here we present an altered form (as described earlier) of the algorithm of [54] incorporating the BTA which switches between GO1 and GO2 when appropriate as stated in §4.3.4. We shall show that our new method provides a much more accurate and consistent (in terms of scattering angle) approximation to the boundary data v with fewer degrees of freedom.

First let us consider the performance of AS1. We take $p = 3$ leading to a total (fixed) number of degrees of freedom of 360 for both angles. Tables 5.6 and 5.7 present results for the incident angles 1 and 2, respectively. The errors shown are for the approximations W_1 and $(\partial u/\partial n)_{go}$ to $\partial u/\partial n$. The relative errors in U_1 were found to be similar to those in W_1 in all cases tested, so we do not reproduce them here. We are aiming for less than 2% error on the boundary with the expectation that this will be reduced in the solution in

120

k_1	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W_1\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-(\frac{\partial u}{\partial \mathbf{n}})_{go}\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W_1\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-(\frac{\partial u}{\partial \mathbf{n}})_{go}\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W_1\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-(\frac{\partial u}{\partial \mathbf{n}})_{go}\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W_1\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-(\frac{\partial u}{\partial \mathbf{n}})_{go}\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W_1\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-(\frac{\partial u}{\partial \mathbf{n}})_{go}\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$
	$n_I = 0.05$		$n_I = 0.025$		$n_I = 0.0125$		$n_I = 0.00625$		$n_I = 0.003125$	
10	1.07e-2	1.11e-1	4.22e-2	1.28e-1	8.23e-2	1.56e-1	1.13e-1	1.83e-1	1.34e-1	2.01e-1
20	7.48e-4	8.03e-2	9.23e-3	8.86e-2	3.77e-2	1.00e-1	7.47e-2	1.22e-1	1.02e-1	1.44e-1
40	1.00e-3	5.87e-2	1.32e-3	6.38e-2	8.11e-3	6.81e-2	3.29e-2	7.64e-2	6.51e-2	9.48e-2
80	5.49e-4	4.18e-2	6.92e-4	4.73e-2	9.40e-4	4.90e-2	6.66e-3	6.10e-2	2.72e-2	5.75e-2
160	7.08e-4	2.91e-2	7.93e-4	3.25e-2	8.93e-4	3.51e-2	1.04e-3	3.67e-2	5.53e-3	3.75e-2

Table 5.6: Performance of approximation space 1 for scattering by a triangle of different refractive index $\mu = 1.5 + n_I \mathrm{i}$ by a plane wave with incident direction $\mathbf{d}^i = (0, -1)$ (Angle 1). Highlighted cells indicate when an error of 2% or better is achieved. The approximation space contains 360 degrees of freedom.

k_1	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W_1\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-(\frac{\partial u}{\partial \mathbf{n}})_{go}\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W_1\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-(\frac{\partial u}{\partial \mathbf{n}})_{go}\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W_1\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-(\frac{\partial u}{\partial \mathbf{n}})_{go}\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W_1\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-(\frac{\partial u}{\partial \mathbf{n}})_{go}\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W_1\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-(\frac{\partial u}{\partial \mathbf{n}})_{go}\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$
	$n_I = 0.05$		$n_I = 0.025$		$n_I = 0.0125$		$n_I = 0.00625$		$n_I = 0.003125$	
10	1.74e-2	1.76e-1	5.73e-2	1.70e-1	9.70e-2	1.73e-1	1.21e-1	1.79e-1	1.33e-1	1.84e-1
20	1.67e-3	1.28e-1	1.05e-2	1.23e-1	4.75e-2	1.21e-1	8.31e-2	1.28e-1	1.06e-1	1.06e-1
40	1.51e-3	9.24e-2	1.75e-3	8.97e-2	1.13e-2	8.63e-2	4.01e-2	8.77e-2	7.08e-2	9.86e-2
80	9.79e-4	6.60e-2	1.02e-3	6.47e-2	1.23e-3	6.31e-2	9.44e-3	6.10e-2	3.38e-2	6.47e-2
160	1.11e-3	4.69e-2	1.15e-3	4.63e-2	1.17e-3	4.56e-2	1.29e-3	4.45e-2	7.96e-3	4.33e-2

Table 5.7: As per Table 5.6 but for an incident wave with direction $\mathbf{d}^i = (\cos(\pi/6), -\sin(\pi/6))$ (Angle 2).
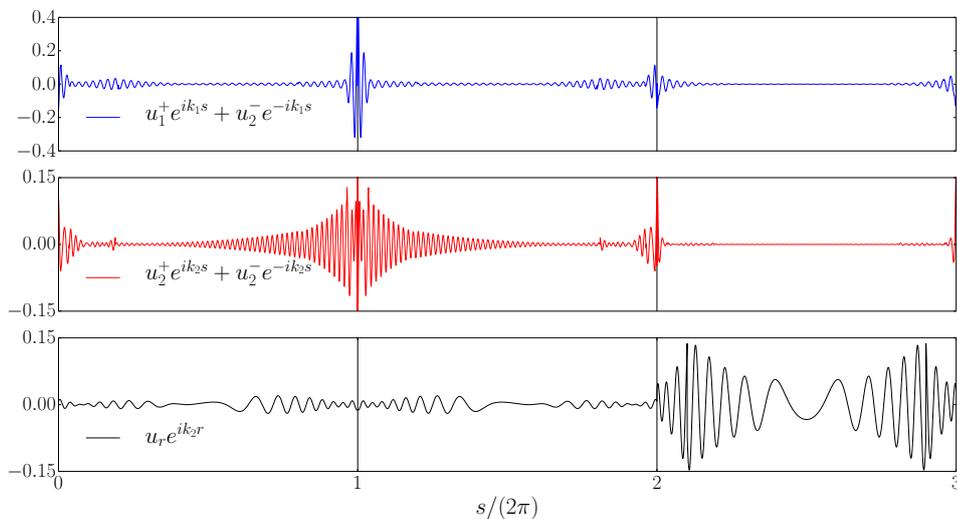
Figure 5.8: Scattering of a plane wave with direction $\mathbf{d}^i = (0, -1)$ (Angle 1) by the triangle in Figure 3.4 . $k_1 = 40$ and $n_I = 0.003125$. The real part of the approximation U_2 split into its oscillatory constituents. Note the different y-axis scales.

the domain and far-field. Therefore, when an error in this range, or better, is achieved, the corresponding cell in the table is highlighted. As can be seen from these two tables, our prescribed accuracy is achieved for $k_1 n_I \geq 0.5$. We emphasise that the GO alone never achieves 2% accuracy for the examples presented.

Now let us assess the efficacy of AS2 for the same examples. This approximation space with $p = 3$ possesses 400 degrees of freedom for Angle 2 and 384 degrees of freedom for Angle 1. Figure 5.8 shows the separate oscillatory constituents of the approximation U_2 for scattering from Angle 2 with $k_1 = 40$ and $n_I = 0.003125$. Clearly the diffraction behaviour on the shadow side (side 3) is dominated by the radial wave with phase $k_2 r$. On the lit sides, the field is dominated by the waves with phases $\pm k_1 s$ and $\pm k_2 s$ with the solution's peaks at the corners being well resolved by the graded meshes. The full results are shown in Tables 5.11 and 5.12.

We see that for Angle 1 the error tolerance is achieved as long as $k_1 n_I \geq 0.125$. However, for Angle 2, the error has scarcely improved from approximation space 1. In order to ascertain the cause for this, let us consider the example $k_1 = 40$ and $n_I = 0.003125$ for Angle 2. The top panel of Figure 5.9 shows a comparison of the real parts of $W_2 - \partial u/\partial \mathbf{n}$ and $\partial u/\partial \mathbf{n} - \partial u_{go}/\partial \mathbf{n}$ on the boundary of the triangle going anti-clockwise starting from $\mathbf{P}_1$. The vertical lines indicate the corners (see Figure 5.10 for a reminder of the scattering setup). That is, we are plotting the best approximation (with AS2) to $\partial u_d/\partial \mathbf{n}$ and the exact $\partial u_d/\partial \mathbf{n}$. The bottom panel shows the absolute error of W_2 around the boundary. It is clear
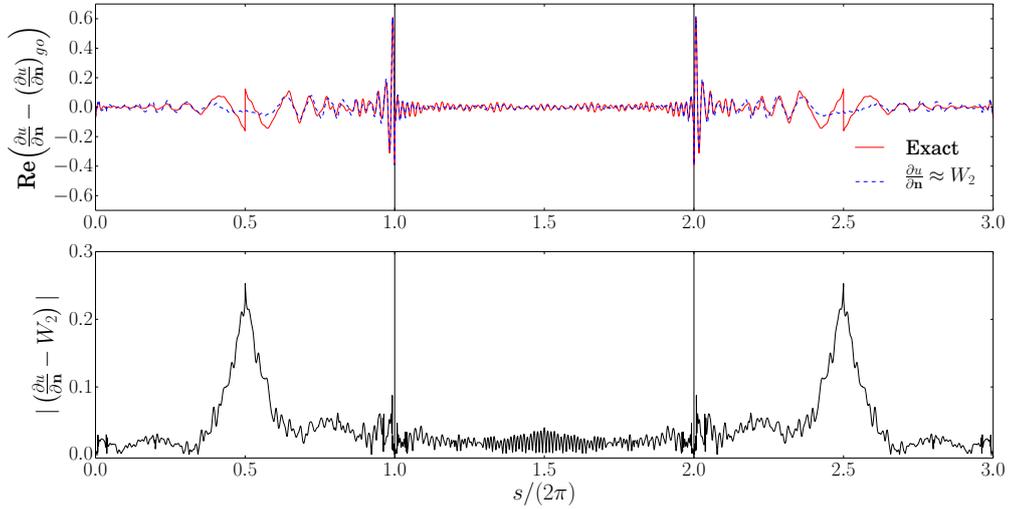
Figure 5.9: Scattering by the triangle in Figure 5.10 with $k_1 = 40$ and $n_I = 0.003125$. The top panel shows the real part of the approximation to $\partial u_d/\partial \mathbf{n}$ with AS2 on Γ compared to the exact value. The bottom panel shows the absolute error in the approximation W_2 on Γ. The error is clearly dominated by the poor approximation near the "second order" shadow boundaries at $s/(2\pi) = 0.5$ and $s/(2\pi) = 2.5$.
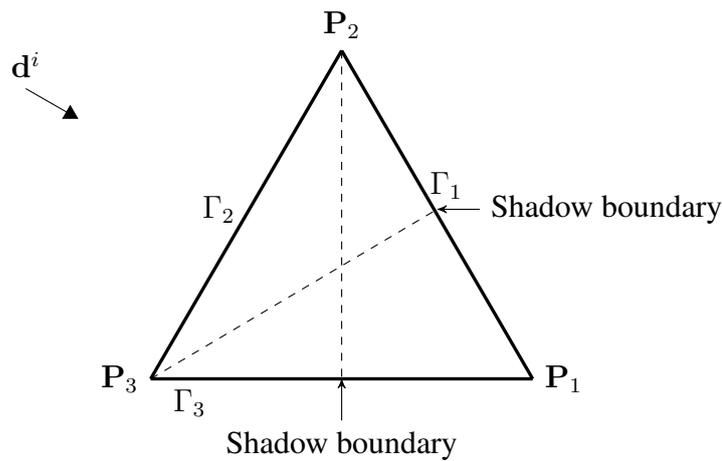
Figure 5.10: Second order shadow boundaries for the triangle in Figure 3.4 for the incident wave associated with Angle 2.

	$n_I = 0.05$		$n_I = 0.025$		$n_I = 0.0125$		$n_I = 0.00625$		$n_I = 0.003125$	
k_1	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_2\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - \left(\frac{\partial u}{\partial \mathbf{n}}\right)_{go}\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_2\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - \left(\frac{\partial u}{\partial \mathbf{n}}\right)_{go}\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_2\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - \left(\frac{\partial u}{\partial \mathbf{n}}\right)_{go}\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_2\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - \left(\frac{\partial u}{\partial \mathbf{n}}\right)_{go}\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_2\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - \left(\frac{\partial u}{\partial \mathbf{n}}\right)_{go}\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$
10	2.94e-3	1.11e-1	7.94e-3	1.28e-1	1.52e-2	1.56e-1	2.29e-2	1.83e-1	2.89e-2	2.01e-1
20	5.20e-4	8.03e-2	2.10e-3	8.86e-2	7.91e-3	1.00e-1	1.70e-2	1.22e-1	2.59e-2	1.44e-1
40	1.00e-3	5.87e-2	9.02e-4	6.38e-2	2.47e-3	6.81e-2	8.80e-3	7.64e-2	1.83e-2	9.48e-2
80	5.48e-4	4.18e-2	6.91e-4	4.73e-2	8.74e-4	4.90e-2	2.06e-3	6.10e-2	8.58e-3	5.75e-2
160	7.07e-4	2.91e-2	7.91e-4	3.25e-2	8.91e-4	3.51e-2	1.00e-3	3.67e-2	2.22e-3	3.75e-2

Table 5.11: Performance of AS2 for scattering from the triangle in Figure 3.4 of refractive index $\mu = 1.5 + n_I$i of a plane wave with incident direction $\mathbf{d}^i = (0, -1)$ (Angle 1). Highlighted cells indicate when an error of 2% or better is achieved. The approximation space contains 400 degrees of freedom.

	$n_I = 0.05$		$n_I = 0.025$		$n_I = 0.0125$		$n_I = 0.00625$		$n_I = 0.003125$	
k_1	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_2\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - \left(\frac{\partial u}{\partial \mathbf{n}}\right)_{go}\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_2\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - \left(\frac{\partial u}{\partial \mathbf{n}}\right)_{go}\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_2\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - \left(\frac{\partial u}{\partial \mathbf{n}}\right)_{go}\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_2\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - \left(\frac{\partial u}{\partial \mathbf{n}}\right)_{go}\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_2\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - \left(\frac{\partial u}{\partial \mathbf{n}}\right)_{go}\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$
10	1.33e-2	1.76e-1	4.20e-2	1.70e-1	7.00e-2	1.73e-1	8.68e-2	1.79e-1	9.52e-2	1.84e-1
20	1.55e-3	1.28e-1	1.08e-2	1.23e-1	3.79e-2	1.21e-1	6.67e-2	1.28e-1	8.52e-2	1.06e-1
40	1.51e-3	9.24e-2	1.70e-3	8.97e-2	9.75e-3	8.63e-2	3.45e-2	8.77e-2	6.12e-2	9.86e-2
80	9.79e-4	6.60e-2	1.02e-3	6.47e-2	1.19e-3	6.31e-2	8.44e-3	6.10e-2	3.03e-2	6.47e-2
160	1.11e-3	4.69e-2	1.15e-3	4.63e-2	1.17e-3	4.56e-2	1.26e-3	4.45e-2	7.34e-3	4.33e-2

Table 5.12: As per Table 5.11 but for an incident wave with direction $\mathbf{d}^i = (\cos(\pi/6), -\sin(\pi/6))$ (Angle 2). Here the approximation space contains 384 degrees of freedom.

that the error in W_2 is dominated by the poor approximation at the points $s/(2\pi) = 0.5$ and $s/(2\pi) = 2.5$. These are the points where Γ is intersected by the "second order" shadow boundaries, associated with the diffracted waves from corners $\mathbf{P}_2$ and $\mathbf{P}_3$, which are currently neglected in our construction of AS2. In the next section we shall show how the error obtained using AS2 can be significantly reduced for configurations such as this one, by a simple modification of the approximation space. This modification will account for second (and potentially higher) order shadow boundaries which introduce discontinuities of a significant magnitude.

5.3.3 Revised Approximation Space 2

Here we develop a revised version of AS2 which performs better than AS2 for low absorptions or low frequencies by taking into account the shadow boundary discontinuities of sufficiently large magnitude introduced by beams of order 2 and higher in the BTA. To do so, we introduce extra points in the meshes for the appropriate "non-adjacent" amplitude functions $v_{2,i+2}$, $i = 1, \ldots, n_s - 2$, as we already do for the lowest order transmitted GO beams in AS2. To limit the complexity of the resulting approximation space, we consider only "strong" discontinuities (as defined below) which make a significant contribution to the error of the AS2 approximation (recall that our target accuracy is approximately 2%). We consider the absolute value of each beam in the BTA which determines the magnitude of these discontinuities. The aim is to determine some tolerance for this absolute value above which we call the discontinuity "strong", and we introduce mesh points corresponding to the beam's shadow boundaries.

The magnitude of the second order beams at the points $s/(2\pi) = 0.5, 2.5$ in Figure 5.9 (for $n_I = 0.003125$, $k_1 = 40$, Angle 2) is 0.488. Comparing the $n_I = 0.003125$ columns of Tables 5.11 and 5.12, we see that we would also like to reduce the error at $k_1 = 80$, for which these beam edges introduce discontinuities with an absolute value of 0.247. So, erring on the side of caution, an initial first tolerance choice might be $\text{tol}_{shad} = 0.2$. However, we must check that such a tolerance does not preclude the inclusion of the first order shadow boundaries for many of the Angle 1 problems. The Angle 1 problems for which we wish to keep the first order shadow boundaries are those for which we obtain a significant improvement from the approximation obtained with AS1 to that obtained with AS2. In particular, those problems for which the error reduces from above 2% in the first instance to below 2% in the second. Comparing tables 5.6 and 5.11, we see that is for $k_1 n_I \leq 0.25$. So we may, without loss of generality, take the case $k_1 = 80, n_I = 0.003125$ for which the first order shadow boundaries are as shown in Figure 5.2. The absolute value of the offending
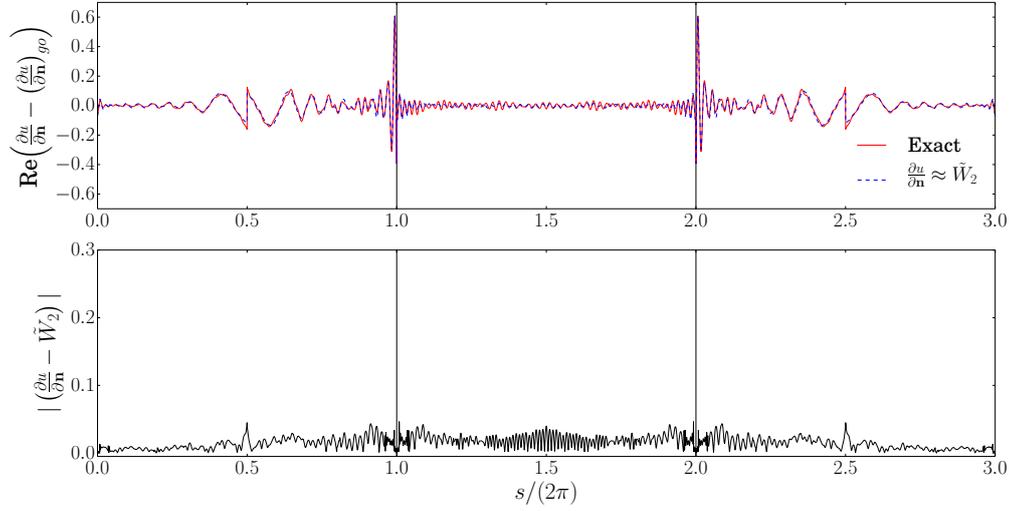
Figure 5.13: As per Figure 5.9 but now using the RAS2 which takes into account the strong shadow boundaries at $s/(2\pi) = 0.5$ and $s/(2\pi) = 2.5$ by introducing mesh points here.

beams at these shadow boundaries is 0.165. Hence we shall decide on a tolerance (again erring on the side of caution) of

$$\text{tol}_{shad} = 0.15. \tag{5.9}$$

Of course, this tolerance ought to be reduced if the user is aiming to obtain a better accuracy in the approximation. However, we shall not explore this further here. This new tolerance defines the "Revised Approximated Space 2" (RAS2) for which we denote the approximation to v as $\tilde{V}_2 = (\tilde{U}_2, \tilde{W}_2)$. Figure 5.13 shows the RAS2 counterpart to the AS2 approximation in Figure 5.9. It is apparent that the approximation $\tilde{W}_2$ is a dramatic improvement over W_2 with the previous error spike of magnitude 2.5 being replaced by a much smaller spike, of magnitude 0.04. To completely eradicate the error spike, geometric grading of the mesh towards the discontinuity is advised (as recommended in [61]), but of course this comes at an increased computational cost. Such a refinement of the technique is not explored here.

Table 5.14 is the RAS2 counterpart to Table 5.12. It can be seen that the required accuracy is achieved when $k_1 n_I \geq 0.125$ as was the case for Angle 1. The number of degrees of freedom is shown also. We see that when tol_{shad} is exceeded by the secondary beams, the number of degrees of freedom increases slightly, from 384 to 400.

k_1	$n_I = 0.05$		$n_I = 0.025$		$n_I = 0.0125$		$n_I = 0.00625$		$n_I = 0.003125$	
	$\dfrac{\|\frac{\partial u}{\partial \mathbf{n}} - \tilde{W}_2\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	#DOF	$\dfrac{\|\frac{\partial u}{\partial \mathbf{n}} - \tilde{W}_2\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	#DOF	$\dfrac{\|\frac{\partial u}{\partial \mathbf{n}} - \tilde{W}_2\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	#DOF	$\dfrac{\|\frac{\partial u}{\partial \mathbf{n}} - \tilde{W}_2\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	#DOF	$\dfrac{\|\frac{\partial u}{\partial \mathbf{n}} - \tilde{W}_2\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	#DOF
10	1.33e-2	384	7.48e-3	400	1.13e-2	400	1.50e-2	400	1.77e-2	400
20	1.55e-3	384	1.08e-2	384	7.52e-3	400	1.62e-2	400	2.50e-2	400
40	1.51e-3	384	1.70e-3	384	9.75e-3	384	6.42e-3	400	1.48e-2	400
80	9.79e-4	384	1.02e-3	384	1.19e-3	384	8.44e-3	384	6.87e-3	400
160	1.11e-3	384	1.15e-3	384	1.17e-3	384	1.26e-3	384	7.34e-3	384

Table 5.14: Performance of RAS2 for scattering from the triangle in Figure 3.4 of different refractive index $\mu = 1.5 + n_I$i of a plane wave from incident Angle 2 ($\pi/6$). Highlighted cells indicate when an error of 2% or better is achieved. The number of degrees of freedom in the approximation space now varies slightly as a function of $k_1 n_I$ but coincides with that for AS2 for sufficiently high $k_1 n_I$.

5.4 Finalised approximation strategy

Tables 5.6, 5.7, 5.11, 5.12 and 5.14 show the accuracy of the various approximation spaces considered and we notice that depending on which $k_1 n_I$ regime we are in, different approximation spaces are necessary to achieve our desired accuracy. In particular, when $k_1 n_I \geq 0.5$ AS1 proves sufficient, whereas for $0.125 \leq k_1 n_I < 0.5$, RAS2 is required to achieve our prescribed accuracy. In this section we shall develop a strategy for switching between the approximation spaces RAS2 and AS1 when the extra DOFs used in RAS2 appear to not be necessary. This strategy shall define the approximation spaces employed in the Galerkin HNA BEM implemented in §6.

We consider the magnitude of the radial wave $v^r \mathrm{e}^{\mathrm{i} k_2 r}$ to investigate when we can switch between the approximation spaces AS1 and RAS2. When this wave is sufficiently small, we expect that neglecting it from our approximation (i.e., reducing from RAS2 to AS1) makes little difference to the accuracy. From the previous chapter, we anticipate that this radially diffracted wave (and similarly for the other diffracted waves) has an amplitude v^r that decays asymptotically as

$$v^r \sim \mathcal{O}\left(\frac{1}{\sqrt{k_2 r}} \mathrm{e}^{-k_1 n_I r}\right), \quad k_1 r \to \infty,$$

where r is the distance from the diffracting corner. Since the exponentially decaying term will dominate as $k_1 n_I r \to \infty$, and also to simplify matters slightly, we consider the amplitude of v^r with the factor $1/\sqrt{k_2 r}$ divided out. That is, we shall identify the decay of v^r with that of the amplitude

$$A := \mathrm{e}^{-k_1 n_I r}. \tag{5.10}$$

For the equilateral triangle of side length 2π, the maximum value of A is

$$\max_r |A(r)| = \mathrm{e}^{-k_1 n_I \pi \sqrt{3}}$$

as can be seen from the geometry of the triangle shown in Figure 5.15. We are interested in the transition between the cases $k_1 n_I = 0.025$ and $k_1 n_I = 0.05$ since this is where the if appears we can switch from RAS2 to AS1 and still retain engineering accuracy. In the first case we wish to keep the $\mathrm{e}^{\mathrm{i} k_2 r}$ functions and in the second case we wish to discard them. The values of $\max_r |A(r)|$ in these two regimes are 0.0658 and 0.257, respectively. We should also keep in mind the tolerance (5.9) for introducing meshpoints at shadow boundaries. This tolerance is 0.15 which fits between the two aforementioned numbers. Again, we decide to err on the side of caution and shall choose our tolerance for discarding the $\mathrm{e}^{\mathrm{i} k_2 r}$-type basis functions as 0.1.
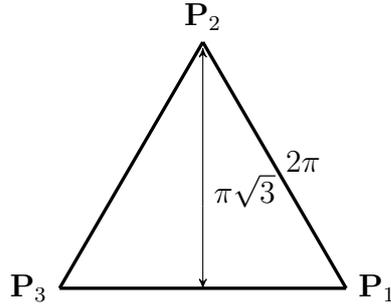
Figure 5.15: Distances within the equilateral triangle.

That is,
$$\text{tol}_{decay} = 0.1. \tag{5.11}$$
So if $\max_r |A(r)| < 0.1$ on an element in the mesh, with $A := \mathrm{e}^{-k_1 N_I r}$, discard the corresponding $\mathrm{e}^{\mathrm{i}k_2 r}$-type basis functions from that element.

If we are to treat the radial waves in this manner, we should also treat the $\mathrm{e}^{\pm \mathrm{i}k_2 s}$ waves in the same way. So,

if $\max_s |A(s)| < \text{tol}_{decay}$ on an element in the mesh, where s is the arc-length from the corner, and
$$A := \mathrm{e}^{-k_1 n_I s},$$
then discard the corresponding $\mathrm{e}^{\mathrm{i}k_2 s}$- or $\mathrm{e}^{\mathrm{i}k_2 s}$-type basis functions from that element.

This proposed approximation space is a hybrid between AS1 and RAS2, and even reduces to a smaller approximation space than AS1 for $k_1 n_I$ large enough. We shall denote the approximation space corresponding to this strategy as AS for "Approximation Strategy". Tables 5.16 and 5.17 show the results from before but using AS. This approximation space optimises the number of degrees of freedom used to achieve the best accuracies from the previous tables. It is of particular note that not only do the number of degrees of freedom not have to increase, they in fact may decrease as k_1 increases to achieve our error tolerance.

	$n_I = 0.05$		$n_I = 0.025$		$n_I = 0.0125$		$n_I = 0.00625$		$n_I = 0.003125$	
k_1	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_3\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	#DOF	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_3\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	#DOF	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_3\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	#DOF	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_3\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	#DOF	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_3\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	#DOF
10	1.07e-2	360	7.94e-3	400	1.52e-2	400	2.29e-2	400	2.42e-2	432
20	7.48e-4	360	9.22e-3	360	7.91e-3	400	1.70e-2	400	2.59e-2	400
40	1.44e-3	312	1.31e-3	360	8.10e-3	360	8.80e-3	400	1.83e-2	400
80	5.61e-4	312	8.69e-4	312	9.38e-4	360	6.65e-3	360	8.58e-3	400
160	7.09e-4	312	7.93e-4	312	9.27e-4	312	1.46e-3	312	5.53e-3	360

Table 5.16: Performance of AS for scattering from the triangle in Figure 3.4 of different refractive index $\mu = 1.5 + n_I \mathrm{i}$ of a plane wave from incident angle 1 ($\pi/2$). Highlighted cells indicate when an error of 2% or better is achieved. The number of degrees of freedom in the approximation space now varies slightly as a function of $k_1 n_I$.

	$n_I = 0.05$		$n_I = 0.025$		$n_I = 0.0125$		$n_I = 0.00625$		$n_I = 0.003125$	
k_1	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_3\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	#DOF	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_3\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	#DOF	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_3\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	#DOF	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - \tilde{W}_2\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	#DOF	$\frac{\|\|\frac{\partial u}{\partial \mathbf{n}} - W_3\|\|}{\|\|\frac{\partial u}{\partial \mathbf{n}}\|\|}$	#DOF
10	1.74e-2	360	7.48e-3	400	1.13e-2	400	1.50e-2	400	1.77e-2	400
20	1.67e-3	360	1.05e-2	360	7.52e-3	400	1.62e-2	400	2.50e-2	400
40	1.97e-3	312	1.75e-3	360	1.13e-2	360	6.42e-3	400	1.48e-2	400
80	1.00e-3	312	1.05e-3	312	1.23e-3	360	9.44e-3	360	6.87e-3	400
160	1.11e-3	312	1.15e-3	312	1.17e-3	312	1.29e-3	312	7.96e-3	360

Table 5.17: As per Table 5.16 but for an incident wave with Angle 2 ($\pi/6$).

5.5 A recap of the algorithm for a general convex polygon

In this section we shall summarise the algorithm developed in this chapter to construct the HNA approximation space for a general convex polygon. The approximation space we have developed relies on certain tolerances for which specific choices have been made here in order to achieve our desired accuracy. These choices can be altered however and so we shall present the algorithm with the tolerances left unspecified (but with reminders of our chosen values given). This is stated for an n_s-sided convex polygon.

Firstly, we calculate the GO term v_{go} using the BTA as described in §4. The BTA is truncated when the amplitude (in absolute value) of the next beam to be included falls below a tolerance tol_{go}. Now, to the approximation of v_d using the HNA method. Recall our ansatz for the unknown v_d on a typical side Γ_j. It is

$$
\begin{aligned}
v_d(\mathbf{x}) \approx & v_{1,1}(s)\mathrm{e}^{\mathrm{i}k_1 s} + v_{1,2}(s)\mathrm{e}^{-\mathrm{i}k_1 s} + v_{2,1}(s)\mathrm{e}^{\mathrm{i}k_2 s} + v_{2,2}(s)\mathrm{e}^{-\mathrm{i}k_2 s} + \\
& + v_{2,3}(r_1)\mathrm{e}^{\mathrm{i}k_2 r_1} + v_{2,4}(r_2)\mathrm{e}^{\mathrm{i}k_2 r_2} + \ldots + v_{2,n_s}(r_{n_s-2})\mathrm{e}^{\mathrm{i}k_2 r_{n_s-2}}.
\end{aligned}
$$

The amplitudes $v_{1,1}, v_{1,2}, v_{2,1}, v_{2,2}$ are approximated by piecewise polynomials supported on overlapping graded meshes. The meshes for $v_{1,1}, v_{2,1}$ are graded geometrically towards $s = 0$ with grading parameters σ_1, σ_2, respectively. The meshes for $v_{1,2}, v_{2,2}$ are graded geometrically towards $s = L_j$ with grading parameters σ_1, σ_2, respectively. These meshes are constructed independently of the incident wave direction. We use one mesh (per side) to support the polynomials approximating the remaining amplitudes $v_{2,3}, v_{2,4}, \ldots, v_{2,n_s}$. The positions of the mesh points in this mesh are dependent on the incident wave direction since they are placed according to the locations of "strong" shadow boundaries introduced by the beam tracing procedure. That is, we first run the BTA and then take note of any beam edges intersecting the side Γ_j. At each of these intersections we evaluate the absolute value of the corresponding beam at the intersection point. If this absolute value is greater than the tolerance tol_{shad}, then we introduce a mesh point at the intersection. This defines the meshes in their initial state.

Now we describe the variation of polynomial degree on the meshes. We choose the polynomial degree vector $\mathbf{p}$ on each of the graded meshes to have the form

$$
(\mathbf{p})_i := \begin{cases} p - \left\lfloor \frac{(n+1-i)}{n}p \right\rfloor, & 1 \leq i \leq n-1, \\ p, & i = n, \end{cases}
$$

where the integer $p \geq 0$ is the highest polynomial degree on the mesh. The number of elements n is related to the maximum polynomial degree p as

$$
n = \lceil C(p+1) \rceil.
$$

On each element of the mesh approximating $v_{2,3}, v_{2,4}, \ldots, v_{2,n_s}$ we place polynomials of the maximum degree p. This defines the degrees of the polynomials accommodated on the meshes.

To avoid potential redundancy (and hence potential ill-conditioning), we employ a process by which we discard basis functions. This only comes into effect for problems where $k_2 \in \mathbb{C}$ so that the functions $\mathrm{e}^{\mathrm{i}k_2 s}, e^{-\mathrm{i}k_2 s}, \mathrm{e}^{\mathrm{i}k_2 r_1}, \mathrm{e}^{\mathrm{i}k_2 r_2}, \ldots, \mathrm{e}^{\mathrm{i}k_2 r_{n_s-2}}$ are decaying. We discard a basis function of the form $P(d)\mathrm{e}^{\mathrm{i}k_2 d}$, for $d = s$ or $d = r$, when

$$\max_d \mathrm{e}^{-\mathrm{i}k_1 n_I d} < tol_{decay},$$

where d ranges over the distances from the diffracting corner to each part of the element supporting the basis function.

We now restate the values of the above tolerances and meshing parameters which were chosen for the implementation of this method here, with the specific aim to achieve a relative error of 2% or less on the boundary Γ. These values are

$$
\begin{aligned}
tol_{go} &= 5 \times 10^{-3}, \\
\sigma_1 &= 0.2, \\
\sigma_2 &= 0.19, \\
C &= 1.5, \\
tol_{shad} &= 0.15, \\
tol_{decay} &= 0.1.
\end{aligned}
\tag{5.12}
$$

In the next chapter we implement this approximation space as a Galerkin BEM. But before doing so we end this chapter by testing its performance for polygons with more than three sides and also investigating its performance as $p \to \infty$.

5.6 Scattering by polygons with more than three sides

The equilateral triangle was used as a prototype scattering shape for the development of the HNA approximation space in this chapter. In this section we shall demonstrate the effectiveness of our HNA approximation space for general convex polygons, by considering scattering by a square with $n = 1.5 + 0.0125\mathrm{i}$ and scattering by a hexagon with $n = 1.5 + 0.00625\mathrm{i}$. The results presented here are for the approximation space of §5.5.

5.6.1 Scattering by a square

We consider the scattering of a plane wave with direction $\mathbf{d}^i = (\cos(\pi/4), -\sin(\pi/4))$ by the square depicted in Figure 5.18 with refractive index $n = 1.5 + 0.0125\mathrm{i}$. We shall
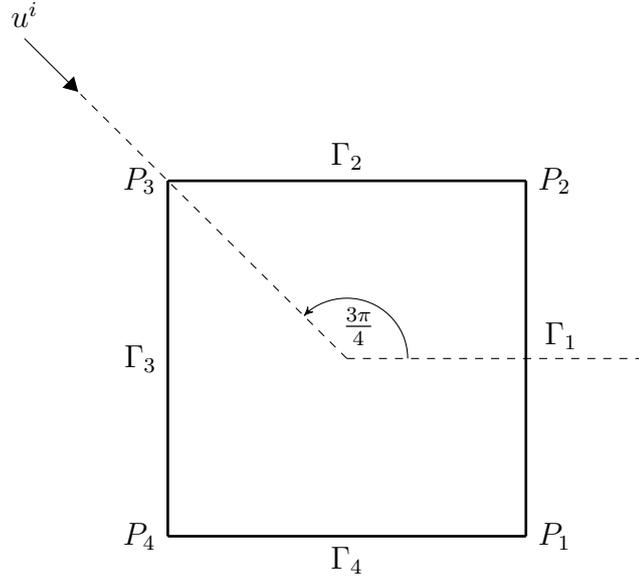
Figure 5.18: Scattering setup for a square of side length 2π.

consider two different values of p. First, we take $p = 3$ as recommended up to now, and then show that, for this example, we can in fact take $p = 2$ and still obtain an accuracy of less than 2% on the boundary for all k_1 considered.

First we present the results for $p = 3$, these are shown in Table 5.19. The results in the

k_1	$\frac{\|\frac{\partial u}{\partial \mathbf{n}} - W_3\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	#DOF	#DOF per λ_2
10	5.29e-3	608	5.07
20	4.34e-3	608	2.53
40	5.01e-3	480	1
80	7.75e-4	480	0.5
160	1.03e-3	416	0.217

Table 5.19: Relative errors in approximation to $\partial u/\partial \mathbf{n}$ on Γ, total number of degrees of freedom in the approximation space and number of degrees of freedom per wavelength (λ_2) for the square using AS with $p = 3$.

table clearly show that the 2% tolerance is easily achieved for all k_1 considered here. Also shown is the number of degrees of freedom per wavelength. This number is calculated as

$$\frac{\#DOF}{2n_s \mathrm{Re}(\mathrm{k}_2)}$$

where here $n_s = 4$ is the number of sides of the polygon. The 2 comes from the fact that there are two unknowns we are approximating (u and $\partial u/\partial \mathbf{n}$) and the product $n_s \mathrm{Re}(\mathrm{k}_2)$ is equal to the number of wavelengths around the boundary since the side lengths of the

133

k_1	$\frac{\|\frac{\partial u}{\partial \mathbf{n}} - W_3\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	#DOF	#DOF per λ_2
10	1.59e-2	448	3.73
20	4.87e-3	448	1.87
40	5.25e-3	352	0.733
80	2.77e-3	352	0.367
160	3.08e-3	304	0.158

Table 5.20: Relative errors in approximation to $\partial u/\partial \mathbf{n}$ on Γ, total number of degrees of freedom in the approximation space and number of degrees of freedom per wavelength (λ_2) for the square using AS with $p = 2$.

square are chosen equal to 2π. The relatively small number in this column of Table 5.19 emphasise the power of the HNA method compared to conventional methods. For the conventional hp BEM presented in §3 it was seen that approximately 3 DOF per λ_2 were required to achieve 1% accuracy in the far-field. The results in the table demonstrate that the HNA method is on a par with the conventional BEM for $k_1 = 10$ but for all higher k_1 it proves much more efficient in terms of memory storage required, with a mere 0.2 DOF per λ_2 required for $k_1 = 160$.

Since the accuracy for $p = 3$ is approximately 0.5% or better on the boundary, we now present in Table 5.20 the results for the AS with $p = 2$. Here we see that again all the boundary errors are less than 2% and so it appears that $p = 2$ is sufficient for this example, to achieve our accuracy goal.

5.6.2 Scattering by a hexagon

Consider the scattering of a plane wave with direction $\mathbf{d}^i = (\cos(\theta^i), -\sin(\theta^i))$, where $\theta^i = \tan^{-1}(2/3)$, by the hexagon in Figure 5.21 with refractive index $n = 1.5 + 0.00625\mathrm{i}$. Table 5.22 shows the results for $p = 3$. We notice an increase in the number of degrees per wavelength compared to the square for $p = 3$. This is due to the increased number of shadow boundaries which lead to additional mesh points. Nevertheless, we observe that less than 2% error is achieved for all wavenumbers with a number of degrees of freedom which is decreasing as k_1 increases.

We have seen that the approximation strategy which we developed in the earlier part of this chapter for the equilateral triangle is also effective for polygons with more sides. In fact, owing to the wider angles of these two latter polygons, the waves must traverse a greater distance inside the shapes, and hence there is greater opportunity for decay. So for these shapes we observe a slight improvement in performance of the HNA approximation space as compared with its performance for the equilateral triangle.
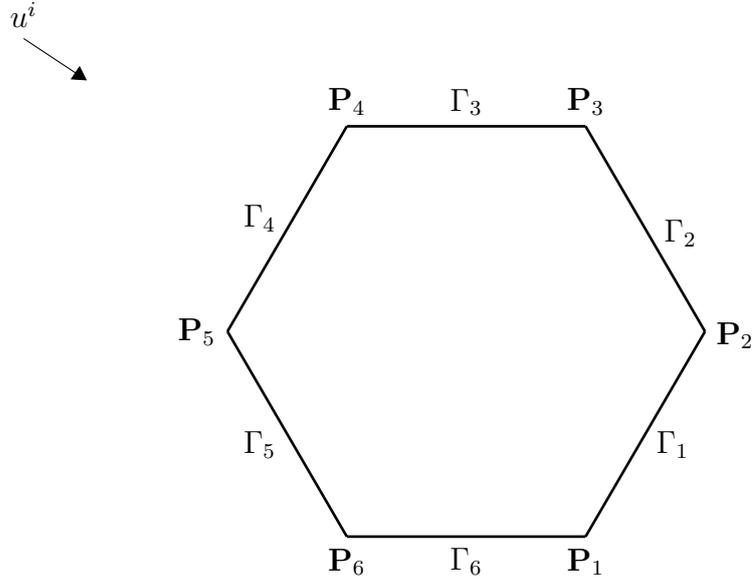
Figure 5.21: Scattering setup for a hexagon of side length 2π.

k_1	$\frac{\lVert \frac{\partial u}{\partial \mathbf{n}} - W_3 \rVert}{\lVert \frac{\partial u}{\partial \mathbf{n}} \rVert}$	#DOF	#DOF per λ_2
10	1.35e-2	1264	7.02
20	1.60e-2	1136	3.16
40	8.06e-3	976	1.36
80	2.00e-3	720	0.5
160	1.21e-3	624	0.217

Table 5.22: Relative errors in approximation to $\partial u/\partial \mathbf{n}$ on Γ, total number of degrees of freedom in the approximation space and number of degrees of freedom per wavelength (λ_2) for the hexagon using AS with $p = 3$.

5.7 Convergence in p

The HNA hp-BEMs developed in [29, 63] (among others) yield errors which decay exponentially with respect to the number of degrees of freedom. Precisely, their approximation φ_N to the solution φ has an error of the form

$$\lVert \varphi - \varphi_p \rVert \leq C \mathrm{e}^{-p\tau}, \tag{5.13}$$

where p is the polynomial degree, and C and τ are constants (see, e.g., [62, 63]) with C growing mildly (algebraically) with k_1. Here $\lVert \cdot \rVert$ represents the L^2 norm on Γ.

These HNA methods are for scattering by impenetrable polygons for which the asymptotic behaviour of the solution is well-understood and hence the approximation spaces can
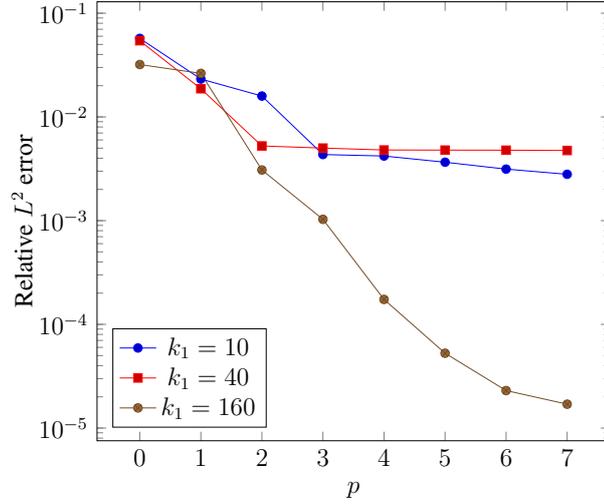
Figure 5.23: Relative L^2 error on Γ versus maximum polynomial degree p for scattering by a square of refractive index $1.5 + 0.0125\mathrm{i}$.

be designed to capture all of the expected oscillatory behaviour. The HNA method proposed here, on the other hand, neglects some of the oscillatory behaviour in the solution and so we cannot anticipate achieving an error bound such as (5.13). Instead we might expect the error to behave like

$$||\varphi - \varphi_p|| \leq C\mathrm{e}^{-p\tau} + F(k_1), \tag{5.14}$$

where

$$F(k_1) = \begin{cases} \mathcal{O}(k_1^{-\alpha}), & \text{if } \mathrm{Im}(k_2) = 0, \\ \mathcal{O}(\mathrm{e}^{-ck_1}k_1^{-\alpha}), & \text{if } \mathrm{Im}(k_2) > 0, \end{cases} \tag{5.15}$$

for some $\alpha > 0$ and $c > 0$. The function F corresponds to the portion of the diffracted field which we are ignoring in the construction of our approximation space. For real k_2 we expect the size of this portion to decay at a rate $k_1^{-\alpha}$, where the constant α is likely to be dependent on the scatterer shape. In contrast, for complex k_2, this remainder term will be exponentially decaying as k_1 increases since now the interior waves all decay exponentially. Therefore, in the complex k_2 case, the error bound (5.13) may be recovered for k_1 sufficiently large.

To explain this claim, we consider the example of scattering of a plane wave with $\mathbf{d}^i = (\cos(\pi/4), -\sin(\pi/4))$ by a square of refractive index $n = 1.5 + 0.0125\mathrm{i}$ for three different wavenumbers $k_1 = 10, 40, 160$. Figure 5.23 shows the relative L^2 error in the approximation $\partial u/\partial \mathbf{n}$ on Γ versus the number of degrees of freedom. It should be noted that each point corresponds to a value of p from 0 to 7.

It can be seen that for $k_1 = 10, 40$, the error reduces quickly to approximately 1×10^{-2} at $p = 3$ but then stagnates and increasing p gives little improvement. This stagnation occurs (which limits our accuracy) since our approximation space ignores higher order asymptotic terms, which are more important for these low values of k_1. For $k_1 = 160$, on the other hand, we appear to achieve exponential convergence in p. The reason for the apparent stagnation at $p = 6, 7$ for $k_1 = 160$ is that we are approaching the accuracy of the reference solution ($\approx 5 \times 10^{-5}$). These results support the form of the error bound (5.14).

Nevertheless, we see for all three wavenumbers that an accuracy of approximately 1% is achieved for $p = 2$. This corresponds to 3.73, 0.733 and 0.158 degrees of freedom per wavelength for $k_1 = 10, 40, 160$, respectively, which is a considerable saving compared to many conventional methods that require between 6 and 10 degrees of freedom per wavelength to achieve engineering accuracy.

Chapter 6

The HNA BEM - an hp Galerkin implementation

In the previous chapter an effective HNA ansatz was developed for a class of transmission problems. This chapter is devoted to the realisation of this approximation space within an hp Galerkin boundary element method.

In §6.1 we state the Galerkin method. In §6.2.1 we discuss some implementation issues associated with the oscillatory basis functions. We conclude the chapter in §6.3 with a selection of numerical examples, demonstrating the performance of the Galerkin method on a range of scattering configurations. The results show that, for the examples considered here, the Galerkin implementation achieves errors very close to the best approximation results from the proceeding chapter. Moreover, all the results presented show that the method produces accurate approximations for all k_1 considered using just a small number of degrees of freedom. Most importantly, the number of degrees of freedom does not need to be increased (in fact, it can be reduced for many problems) in order to maintain accuracy as k_1 increases.

6.1 Galerkin method

Let us begin by recalling from §2 the integral equation we wish to solve, however with one small difference: here we shall solve for the unknowns scaled as $\left(u, \frac{1}{k_1} \frac{\partial u}{\partial \mathbf{n}}\right)$ rather than $\left(u, \frac{\partial u}{\partial \mathbf{n}}\right)$ as was done in §2. This is a natural scaling since we expect the absolute value of $\partial u / \partial \mathbf{n}$ to be of the order of k_1. So, the integral equations we wish to solve are the following. Given an incident field

$$v^i = \frac{1}{k_1} \begin{pmatrix} u^i \\ \alpha \frac{\partial u^i}{\partial \mathbf{n}} \end{pmatrix} \in L^2(\Gamma), \tag{6.1}$$

we seek the solution

$$v = \begin{pmatrix} u \\ \frac{1}{k_1} \frac{\partial u}{\partial \mathbf{n}} \end{pmatrix} \in L^2(\Gamma) \tag{6.2}$$

such that

$$\mathcal{A}v = v^i, \tag{6.3}$$

where

$$\mathcal{A} = \begin{pmatrix} \frac{1}{2k_1}(1+\alpha)I + \frac{1}{k_1}(\alpha\mathcal{D}_2 - \mathcal{D}_1) & \mathcal{S}_1 - \mathcal{S}_2 \\ \frac{1}{k_1}(\mathcal{H}_1 - \mathcal{H}_2) & \frac{1}{2}\left(1 + \frac{1}{\alpha}\right)I + \left(\mathcal{D}_1' - \frac{1}{\alpha}\mathcal{D}_2'\right) \end{pmatrix}. \tag{6.4}$$

Here, as in previous chapters, we shall take $\alpha = 1$. The HNA method proposed here should also be applicable to the $\alpha \neq 1$ case, however, an investigation into this claim is left to future work.

In the HNA approach, we decompose the solution as

$$v = v_{go} + v_d, \tag{6.5}$$

where v_{go} is the GO approximation calculated using the BTA (as described in §4.2), and v_d is the diffracted field which we approximate using our HNA BEM. More explicitly, the decomposition (6.5) is substituted into (6.3) to give

$$\mathcal{A}v_d = v_i - \mathcal{A}v_{go} \tag{6.6}$$

which is the form of equation (6.3) we aim to solve.

Having, in the previous chapter, designed an approximation space Q_{N,k_1} for v_d, we now seek an element of Q_{N,k_1} using the Galerkin method. That is, we seek $V = (U, W) \in Q_{N,k_1} \subset L^2(\Gamma)$ such that

$$\langle \mathcal{A}V, q \rangle_{L^2(\Gamma)} = \langle v^i - \mathcal{A}v_{go}, q \rangle_{L^2(\Gamma)}, \quad \text{for all } q \in Q_{N,k_1}. \tag{6.7}$$

One of the desirable properties of the Galerkin method is that is can often be proven to yield quasi-optimal approximations (as defined below). Such quasi-optimality relies upon the coercivity of the boundary integral operator.

For the problem of scattering by a sound-soft polygon (see [63]), there exists an integral equation formulation which is provably coercive for star-shaped Lipschitz scatterers, namely the "star-combined" formulation [127]. The definition of a coercive operator is as follows. The boundary integral operator $\mathcal{A}$ is defined as coercive if, for some $\gamma > 0$ (the coercivity constant),

$$|\langle \mathcal{A}v, v \rangle| \geq \gamma ||v||^2, \quad \forall v \in L^2(\Gamma). \tag{6.8}$$

139

(We note that the more general definition of coercivity extends to general Hilbert spaces, however we focus on the space $L^2(\Gamma)$ here since it is of primary interest.) The coercivity and continuity of the integral operator lead, via Céa's lemma to the quasi-optimality estimate

$$||v_d - V|| \leq \frac{C(k_1)}{\gamma} \inf_{q \in Q_{N,k_1}} ||v_d - q||, \tag{6.9}$$

where C is a constant which may depend on k_1.

For many previous HNA methods (e.g., [63] for the impenetrable convex polygon, [62] for the 2D screen), it is assumed that the corresponding integral operator is coercive (e.g., [63, Assumption 2.3]), and since the operator is also continuous, Céa's lemma implies the quasi-optimality estimate

$$||v_d - V||_{L^2(\Gamma)} \leq C(k_1) \inf_{q \in Q_{N,k_1}} ||v_d - q||_{L^2(\Gamma)}. \tag{6.10}$$

The "standard combined" integral equation formulation for the sound-soft problem has also been proven in [129] to be coercive for Γ sufficiently smooth, and numerical evidence in [14] suggests that it is true for general polygons also.

In the penetrable case, however, there is little known about the coercivity or otherwise of boundary integral operator (6.4). In fact, we are not aware of a coercive boundary integral equation formulation for this problem. Therefore, a wavenumber explicit quasi-optimality result such as (6.10) is not available. Nevertheless, it is worthwhile to investigate the relationship between the Galerkin approximation and the best approximation since we hope that the Galerkin method can generate approximations from our designed approximation space which are close to optimal. In the results presented later, we shall calculate the ratio

$$C_{QO} := \frac{||\frac{\partial u_d}{\partial \mathbf{n}} - W||}{||\frac{\partial u_d}{\partial \mathbf{n}} - W_{BA}||} \tag{6.11}$$

where W_{BA} is the best approximation to $\partial u / \partial \mathbf{n}$ obtained using the least squares method of the previous chapter. We call the quantity (6.11) the "quasi-optimality ratio". The results presented later in this chapter suggest that the Galerkin approximation may indeed be quasi-optimal with a constant C_{QO} that is only mildly dependent on k_1 and the scattering geometry. However, more in-depth experiments would be required to say anything more definitive.

6.2   Implementation

Much of the implementation is shared with that of the conventional BEM presented in §3. The only differences arise in the numerical quadrature and in the scaling of the basis

functions. The integrals to be evaluated are now more complicated owing to the many oscillations in the kernels across the supports of the basis functions. Hence we must modify the numerical quadrature to subdivide the integration domain according to the wavelength of the oscillations, and on each of these subdomains employ the quadrature techniques described in §3.2.3. This is potentially expensive and as a cost which scales as k_1^2 (for the Galerkin method), however there exist more sophisticated method for such integrals which have a cost independent of k_1. Such methods are not implemented here, however we give a brief review of them in §6.2.2. The scaling of the basis functions is necessary since, for complex k_2, the absolute value of the basis functions varies exponentially across their supports. So this can lead to entries in the system matrix which vary enormously which may cause ill-conditioning of the system. This is different to previous HNA methods in which the wavenumber is real.

6.2.1 Normalising the basis functions

In §3.2.1 we reviewed how the Legendre polynomials can be normalised in order to form an orthonormal basis. Such a scaling is beneficial since it is desirable for numerical reasons (in particular, conditioning) to have a system matrix with entries of a similar magnitude, usually between 0 and 1. The basis functions in our HNA approximation space are not orthogonal, however it is still beneficial to normalise them appropriately. Consider the oscillatory basis function

$$\rho(s) := L_j(s) \exp\{\mathrm{i}k_1\psi(s)\}, \tag{6.12}$$

where $L_j(s)$ is the Legendre polynomial of order j as defined in §3.2.1 and with support $[a, b]$, and $\psi(s)$ is one of the phases chosen in the previous chapter. Then the local L^2 norm of $\rho(s)$ is

$$||\rho||_{L^2[a,b]} = \left[\int_a^b L_j(s) \exp\{\mathrm{i}k_1\psi(s)\} \overline{L_j(s) \exp\{\mathrm{i}k_1\psi(s)\}} \mathrm{d}s \right]^{\frac{1}{2}}$$

$$= \left[\int_a^b (L_j(s))^2 \exp\{-2k_1 \mathrm{Im}[\psi(\mathrm{s})]\} \mathrm{d}s \right]^{\frac{1}{2}}$$

It is easy to see that, if $\psi(s)$ is real valued, then $||\rho|| = \sqrt{(b-a)/(2j+1)}$ as we saw for the Legendre polynomials on their own in §3.2.1. However, if $\psi(s)$ is complex-valued, as is the case for the functions oscillating with wavenumber k_2 when $k_2 \in \mathbb{C}$, we may have that $||\rho||$ is extremely large or extremely small. To see this more clearly, consider the example $\psi(s) = k_2 s = k_1(n_R + n_I \mathrm{i})$. Then we have that

$$||\rho||_{L^2[a,b]} = \frac{1}{\sqrt{2k_1 n_I}} \left(\mathrm{e}^{-2k_1 n_I a} - \mathrm{e}^{-2k_1 n_I b} \right)^{1/2}$$

which tends to zero as $k_1 \to \infty$. In a similar way, the phase $\rho(s) = -k_2 s$ leads to an exponentially growing value for $||\rho||$. Therefore, in our implementation we normalise each basis function by its local L^2 norm in order to avoid very large and very small entries occurring in the system matrix.

6.2.2 Oscillatory integrals

In calculating the system matrix for the Galerkin HNA BEM we are required to evaluate integrals of the form

$$I := \int_{\Gamma_j} \int_{\Gamma_i} k(\mathbf{x}, \mathbf{y}) \phi_i(\mathbf{x}) \overline{\phi_j(\mathbf{y})} \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} \qquad (6.13)$$

where $k(\mathbf{x}, \mathbf{y})$ is a kernel function composed of a difference of (oscillatory) Hankel functions (see equation (3.14) for details), and ϕ_i, ϕ_i are oscillatory basis functions from the HNA approximation space constructed in §5. Γ_i and Γ_j are the supports of ϕ_i and ϕ_j, respectively. When the wavenumber k_1 is large with respect to the supports Γ_j and Γ_j, integrals of the form (6.13) are highly oscillatory.

To achieve an efficient implementation, one should apply oscillatory quadrature techniques such as Filon quadrature, Levin collocation and Numerical Steepest Descent (NSD). Such methods require a number of quadrature points that remains fixed, or even decreases, as k_1 increases. Applications of the Filon method to integrals of the form (6.13) have been reported in [81] for smooth scatterers, and in [135] for polygonal scatterers. In [67] the NSD was presented for simple examples of 2D oscillatory integrals, and its application to the specific type of 2D integrals arising in Galerkin HNA BEM for 2D polygonal scatterers was briefly discussed in [70]. For detailed reviews of oscillatory quadrature techniques, the reader is referred to [69, 72].

The focus of this thesis, however, is the design of an HNA approximation space for the transmission problem and not a fast implementation and hence here we employ a brute force approach to the evaluation of the oscillatory integrals. This brute force technique is to divide the integration domain into a large number of smaller domains over which the integrand is not oscillatory, then apply standard quadrature rules (such as those discussed in §3.2.3) on the integrals over each of the subdomains. This process is simple and robust but expensive; the number of quadrature points required scales with k_1^2. This technique was employed to construct the system matrix for the results presented in §6.3 owing to the ease of implementation. We now describe this technique is more detail.

We begin with the case when the two basis functions are well separated so that the singularity in the kernel is sufficiently far away (as defined in §3.2.3). Define the integers
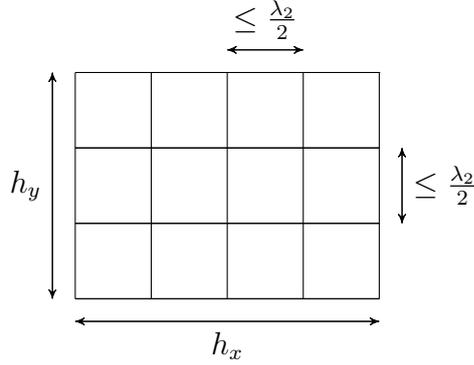
Figure 6.1: Composite Gaussian quadrature for an oscillatory integral.

n_x and n_y as

$$n_x = \left\lceil \frac{h_x}{\lambda_2} \right\rceil \quad \text{and} \quad n_y = \left\lceil \frac{h_y}{\lambda_2} \right\rceil. \tag{6.14}$$

Then divide the integration region shown in Figure 6.1 of area $h_x \times h_y$ into $n_x n_y$ boxes of size $\frac{h_x}{n_x} \times \frac{h_y}{n_y}$. Thus none of the smaller boxes has a dimension greater than $\frac{\lambda_2}{2}$. This ensures that the functions we are integrating have no more than half an oscillation in either of the x and y directions, hence employing a tensor product of two Gauss-Legendre quadrature rules will provide a reliable and accurate approximation on each of these small boxes. It was found here that a number of quadrature points in either direction of

$$n_G = 10 + p,$$

where p is the polynomial degree, is sufficient to achieve a relative error of better than 10^{-14} for the integration.

If the two basis functions touch at one end then we have a logarithmic singularity located at one corner of the integration domain as depicted by the red point in Figure 6.2(b). As suggested in §3.2.3.3 we may use the tensor product of two one-dimensional composite Gauss rules to take account of the singularity. However, this does not take account of the highly oscillatory nature of the integrand. Therefore we modify this method suggested in §3.2.3.3 in the following way.

First we subdivide the integration domain using the geometric grading as described in §3.2.3, with grading parameter $\sigma = 0.15$. This results in a grid as depicted in Figure 6.2(a). Next, we subdivide any boxes which have a dimension larger than $\lambda_2/2$ as shown in Figure 6.2(b). This is done in the same fashion as described above for the non-singular case. Then on each of the boxes we use $10 + p$ Gaussian quadrature points in either direction.

Finally, for the case when the supports of the two basis functions coincide we have a singularity along the diagonal of the integration domain. To cope with this case, we first

143

(a) First step: grade towards the singularity with grading parameter $\sigma = 0.15$.

(b) Second step: subdivide the boxes into smaller boxes of dimension smaller than or equal to $\lambda_2/2$.
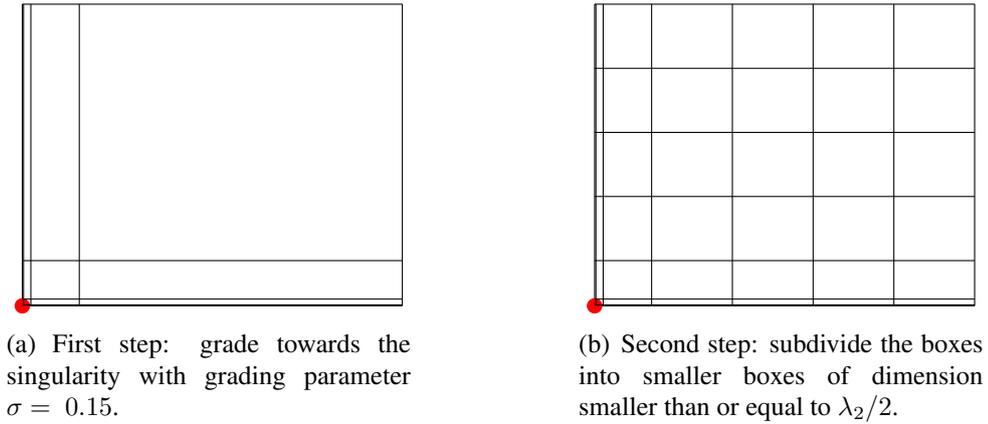
Figure 6.2: Composite quadrature for an oscillatory two-dimensional integral with a logarithmic singularity at one corner of the integration domain.

split the domain along this diagonal into two triangular regions as was done in §3.2.3.3. We also follow the approach there of employing the Duffy transformation on each of these triangles which maps them to squares on which the integrands have singularities along two of the sides of the domain. Now the integrals over each of these two transformed domains may be treated in the same way as discussed for the singularity at a corner above.

6.3 Numerical results

In this section we present numerical results demonstrating the performance of our HNA Galerkin BEM for a range of scattering problem involving triangle, squares and hexagons. Recall that throughout we are taking $\alpha = 1$.

6.3.1 Scattering by a triangle

We first consider scattering by the triangle of Figure 3.4. We consider the following six combinations of incident wave direction and refractive index:

1. $\mathbf{d}^i = (0, -1)$ and $n = 1.5 + 0.05\mathrm{i}$,

2. (i) $\mathbf{d}^i = (0, -1)$ and $n = 1.5 + 0.0125\mathrm{i}$,

   (ii) $\mathbf{d}^i = (\cos(\pi/3), -\sin(\pi/3))$ and $n = 1.5 + 0.0125\mathrm{i}$,

   (iii) $\mathbf{d}^i = (\cos(\pi/6), -\sin(\pi/6))$ and $n = 1.5 + 0.0125\mathrm{i}$,

3. $\mathbf{d}^i = (0, -1)$ and $n = 1.5 + 0.00625\mathrm{i}$,

4. $\mathbf{d}^i = (0, -1)$ and $n = 1.5 + 0.003125\mathrm{i}$,

144

Five of these examples were considered in §5 and are considered again for direct comparison to the best approximation results. The one additional example is 2.(ii) where a new incident direction is considered, namely $\mathbf{d}^i = (\cos(\pi/3), -\sin(\pi/3))$. This direction leads to a grazing incident wave and is an intermediate direction between the two previously considered directions of §5. For grazing incidence we expect the head wave discussed in §4.8.1 to be at its most prominent since it is generated by the incident wave as well as the diffracted waves, rather than just the diffracted waves for non-grazing incidence. Recall that head waves are neglected in our approximation space, therefore it is interesting to see how our HNA BEM fares for this grazing incidence case.

In addition to these examples, we shall present a comparison of the accuracies of the approximations obtained for scattering by a triangle with $n_I = 0, 1/320, 1/160, 1/80, 1/40, 1/20$ by a wave with $\mathbf{d}^i = (0, -1)$. This is to show how the accuracy improves with increasing absorption but also to demonstrate that the method produces an approximation satisfactory for many applications even for scatterers with zero absorption. In all of our experiments in this subsection we take the maximum polynomial degree as $p = 3$.

$$\textbf{Example 1: } \mathbf{d}^i = (0, -1) \textbf{ and } n = 1.5 + 0.05\mathrm{i}.$$

The scattering setup is as depicted in Figure 3.4 with the incident wave travelling at an angle of $\pi/2$ to the x-axis (Angle 1). Relative L^2 errors demonstrating the accuracy of the approximation on the boundary obtained by the Galerkin method are presented in Table 6.3 and should be compared with the best approximation results in Table 5.16. Also given are the relative L^2 errors for the Galerkin approximation to the far-field F_G, the total field u_G on the circle shown in Figure 5.5, the condition numbers of the system matrix A, the number of degrees of freedom per wavelength in the approximation space, and the quasi-optimality ratio C_{QO}. We see that the error in the approximation on the boundary is below

k_1	$\frac{\|\frac{\partial u}{\partial \mathbf{n}} - W\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|F - F_G\|}{\|F\|}$	$\frac{\|u - u_G\|}{\|u\|}$	cond(A)	# DOF	# DOF per λ_2	C_{QO}
10	1.07e-2	9.66e-3	5.06e-3	5.66e5	360	4	1.01
20	7.92e-4	7.16e-4	3.82e-4	6.91e5	360	2	1.07
40	1.66e-3	6.65e-4	3.58e-4	3.80e5	312	0.867	1.15
80	6.70e-4	3.18e-4	1.68e-4	2.25e5	312	0.433	1.19
160	1.14e-3	7.32e-4	4.01e-4	1.50e6	312	0.217	1.61

Table 6.3: Scattering by a triangle with refractive index $n = 1.5 + 0.05\mathrm{i}$ by plane wave with direction $\mathbf{d}^i = (0, -1)$. Relative L^2 errors, condition numbers, and the number of DOF per wavelength in the approximation space.

2% for all k_1, and the errors in the far-field and the total field are less than 1% for all k_1. We

also observe that the conditioning of the system matrix has a value of approximately 10^5 or 10^6 for each k_1. This is an admissible condition number since we have ensured that the entries of the matrix are calculated to an accuracy of nearly 10^{-14} relative error. It should be noted that this would not be the case if the basis function discarding process of §5.4 and the basis function scaling of §6.2.1 were not incorporated. Results demonstrating this are not presented here, however these problems were indeed encountered in the development of this method.

Also shown is the number of degrees of freedom per wavelength. These results demonstrate that the HNA method is on a par with the conventional BEM for $k_1 = 10$ but for all higher k_1 it proves much more efficient in terms of memory storage required, with a mere 0.2 DOF per λ_2 required for $k_1 = 160$. It should also be noted that for $k_1 \geq 20$ the error in the far-field is less than 0.1%.

Finally, the table shows the values of the quasi-optimality ratio. For low k_1, C_{QO} has a value close to 1 which increases slowly with increasing k_1. For all k_1 considered however, C_{QO} does not exceed 2.

<div align="center">Example 2: $n = 1.5 + 0.0125\mathrm{i}$.</div>

(i) $\mathbf{d}^i = (0, -1)$
(ii) $\mathbf{d}^i = (\cos(\pi/3), -\sin(\pi/3), -1)$
(iii) $\mathbf{d}^i = (\cos(\pi/6), -\sin(\pi/6), -1)$

The results for these three examples are presented in Table 6.4, Table 6.5 and Table 6.6, respectively. We again see that the Galerkin method produces approximations close to the best approximation results with a quasi-optimality ratio close to 1 but increasing slightly with k_1. For examples (i) and (ii), we observe that the error on the boundary is less than 2% for all k_1. Also, we see that the approximation to the total field u in the domain is even better.

For (ii) however, we observe that the error on the boundary is not less than 2% until k_1 exceeds 40. We conjecture that this large error is due to the prominence of the head wave in this grazing incidence case. To see the effect of this head wave, we plot in Figure 6.7 the exact $(\partial u/\partial n - (\partial u/\partial n)_{go})$ with its approximation via the HNA BEM in the top panel, along with the absolute value of the difference between these two in the lower panel, i.e., the absolute error in W on Γ. We see that the most major contribution to the error is coming from the approximation on the third side, Γ_3, of the triangle $(2 \leq s/(2\pi) \leq 3)$ which is precisely where we expect the head wave to be most prominent. Recall that the incident wave grazes the first side, Γ_1, therefore the head wave it generates propagates into the triangle and impinges on Γ_3 (see Figure 3.4 for a reminder of the side labelling). The error

has a regular oscillatory form which strengthens the conjecture that this is the head wave we are observing.

k_1	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|F-F_G\|}{\|F\|}$	$\frac{\|u-u_G\|}{\|u\|}$	cond(A)	# DOF	# DOF per λ_2	C_{QO}
10	1.55e-2	1.25e-2	7.01e-3	7.24e5	400	4.44	1.07
20	8.31e-3	7.51e-3	4.19e-3	6.02e5	400	2.22	1.07
40	9.15e-3	5.76e-3	3.07e-3	3.74e5	360	1	1.18
80	1.23e-3	8.31e-4	4.52e-4	2.60e5	360	0.5	1.31
160	1.26e-3	2.35e-3	1.31e-3	2.14e5	312	0.217	1.39

Table 6.4: Scattering by a triangle with refractive index $n = 1.5 + 0.0125i$ by plane wave with direction $\mathbf{d}^i = (0, -1)$. Relative L^2 errors, condition numbers, and the number of DOF per wavelength in the approximation space.

k_1	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|F-F_G\|}{\|F\|}$	$\frac{\|u-u_G\|}{\|u\|}$	cond(A)	# DOF	# DOF per λ_2	C_{QO}
10	6.80e-2	7.36e-2	4.13e-2	7.80e5	400	4.44	1.02
20	4.38e-2	4.38e-2	2.14e-2	5.92e5	392	2.18	1.04
40	1.99e-2	2.06e-2	1.05e-2	3.74e5	360	1	1.16
80	7.38e-3	8.91e-3	4.55e-3	2.60e5	360	0.5	1.31
160	4.05e-3	6.36e-3	3.23e-3	2.14e5	312	0.217	1.40

Table 6.5: Scattering by a triangle with refractive index $n = 1.5 + 0.0125i$ by plane wave with direction $\mathbf{d}^i = (\cos(\pi/3), -\sin(\pi/3))$. Relative L^2 errors, condition numbers, and the number of DOF per wavelength in the approximation space.

k_1	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|F-F_G\|}{\|F\|}$	$\frac{\|u-u_G\|}{\|u\|}$	cond(A)	# DOF	# DOF per λ_2	C_{QO}
10	1.17e-2	1.46e-2	7.68e-3	7.93e5	400	4.44	1.06
20	7.67e-3	9.14e-3	4.67e-3	6.43e5	400	2.22	1.02
40	1.14e-2	1.41e-2	6.85e-3	3.74e5	360	1	1.17
80	1.18e-3	1.06e-3	5.33e-4	2.60e5	360	0.5	1.33
160	1.47e-3	1.74e-3	9.25e-4	2.14e5	312	0.217	1.40

Table 6.6: Scattering by a triangle with refractive index $n = 1.5 + 0.0125i$ by plane wave with direction $\mathbf{d}^i = (\cos(\pi/6), -\sin(\pi/6))$. Relative L^2 errors, condition numbers, and the number of DOF per wavelength in the approximation space.

Also we observe that the condition numbers are well-behaved and in fact smaller than for the previous example which has stronger absorption. Here the condition numbers are less than 10^6 for all k_1. Finally we note that the value of C_{QO} does not vary much with
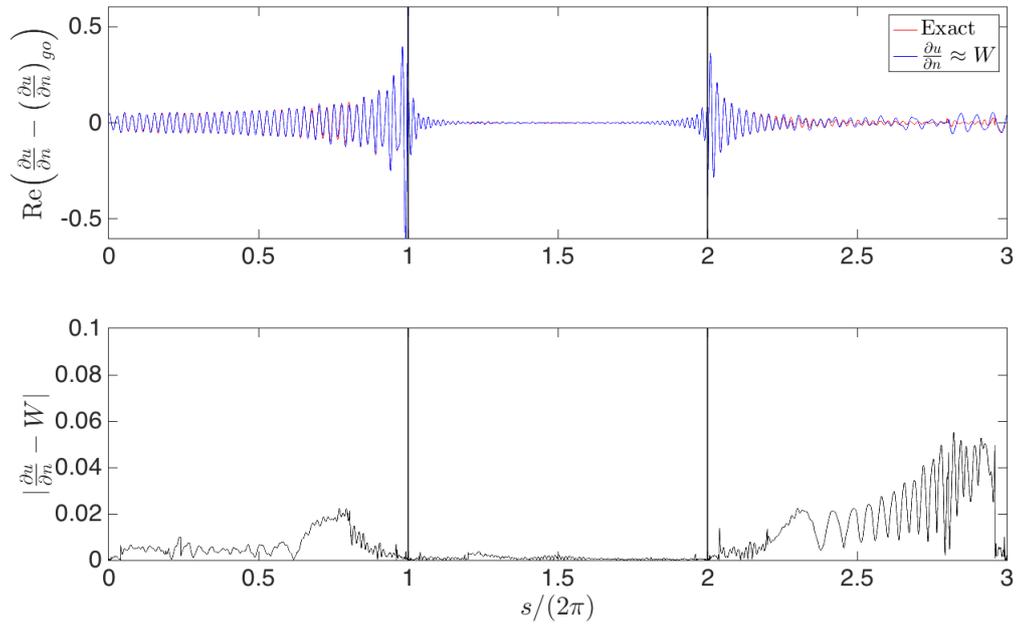
Figure 6.7: Scattering by the triangle in Figure 6.8 with $k_1 = 40$ and $n = 1.5 + 0.0125i$ for grazing incidence. The top panel shows the real part of the approximation to $(\partial u/\partial \mathbf{n} - (\partial u/\partial \mathbf{n})_{go})$ with the HNA BEM on Γ compared to its exact value. The bottom panel shows the absolute error in the approximation W on Γ.
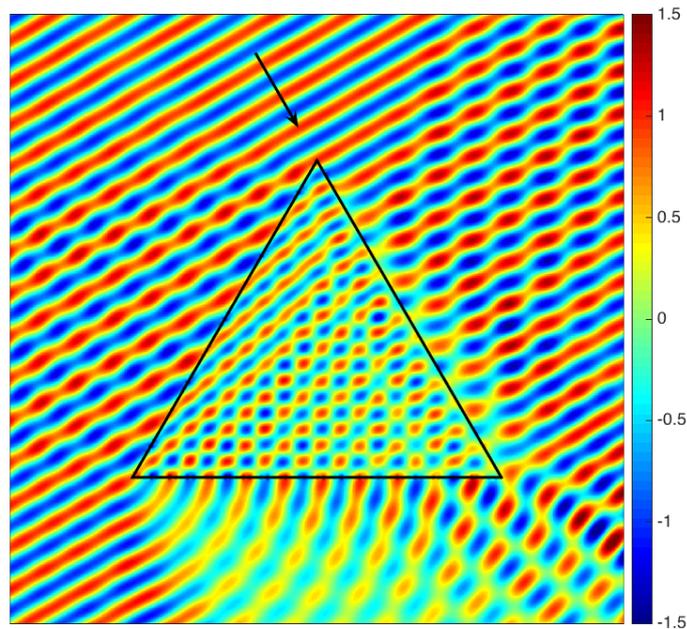
Figure 6.8: Scattering of a plane wave at grazing incidence $\mathbf{d}^i = (\cos(\pi/3), -\sin(\pi/3))$ with $k_1 = 10$ an equilateral triangle of refractive index $n = 1.5 + 0.0125i$.

148

incident direction. This is to perhaps be expected since the system matrix remains largely the same for all incident directions. In fact, for this example, for $k_1 \geq 40$, the matrix is exactly the same for all three incident directions.

Example 3: $\mathbf{d}^i = (0, -1)$ **and** $n = 1.5 + 0.00625\mathrm{i}$**.**

For this example, the absorption is again reduced. Nevertheless, we achieve an approximation of less than 2% on the boundary for all $k_1 \geq 20$ and the error is decreasing with increasing k_1 despite the fact that the number of degrees of freedom per wavelength is decreasing. The errors in the far-field are also decreasing with increasing k_1 and are always less than 2%. The errors in the total field on the circle in Figure 5.5 are always less than 1% and are decreasing with increasing k_1. An example of the total field on this circle is shown in Figure 6.10 for $k_1 = 80$ where ϑ runs from 0 to 2π anti-clockwise starting at the positive x-axis. We can clearly see the strong reflection regions at $\vartheta/2\pi = 0.25$ and 0.75, and the shadow region behind the scatterer at $\vartheta/2\pi = 0.5$.

k_1	$\frac{\|\frac{\partial u}{\partial \mathbf{n}} - W\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|F - F_G\|}{\|F\|}$	$\frac{\|u - u_G\|}{\|u\|}$	cond(A)	# DOF	# DOF per λ_2	C_{QO}
10	2.36e-2	1.96e-2	9.43e-3	3.92e6	400	4.44	1.03
20	1.76e-2	1.57e-3	7.49e-3	8.98e5	400	2.22	1.04
40	8.98e-3	8.95e-3	4.15e-3	4.76e5	400	1.11	1.02
80	6.68e-3	6.53e-3	2.83e-3	2.09e5	360	0.5	1.00
160	1.87e-3	1.06e-3	4.85e-4	3.80e5	312	0.217	1.28

Table 6.9: Scattering by a triangle with refractive index $n = 1.5 + 0.00625\mathrm{i}$ by plane wave with direction $\mathbf{d}^i = (0, -1)$. Relative L^2 errors, condition numbers, and the number of DOF per wavelength in the approximation space.

The condition numbers are again well-behaved and are around 10^5 or 10^6. Finally, we see that the value of C_{QO} is close to one and starts to increase at $k_1 = 160$. C_{QO} is smaller for this example than the previous examples considered. This is perhaps due to the fact that the singularities at the corners of the triangle are weaker for scatterers with lower absorption, and it is the strengthening of these singularities that is leading to the increase in C_{QO}. We can test this hypothesis with the next example.

Example 4: $\mathbf{d}^i = (0, -1)$ **and** $n = 1.5 + 0.003125\mathrm{i}$**.**

This example reduces the absorption further and we see, as conjectured in the previous example, that indeed the quasi-optimality ratio appears to become better behaved as the
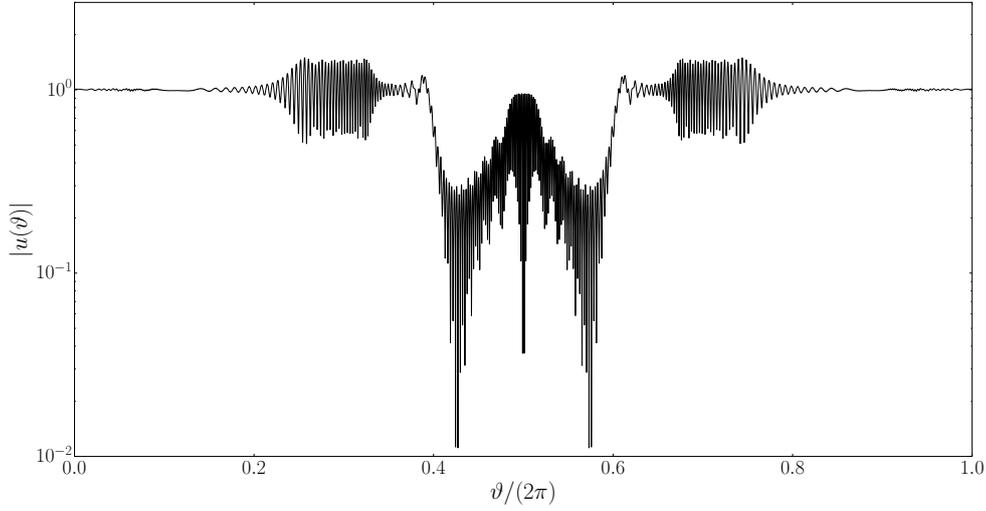
Figure 6.10: Total field on the circle in Figure 5.5 pattern for $k_1 = 80$. $n = 1.5 + 0.00625\mathrm{i}$. Incident direction $\mathbf{d}^i = (0, -1)$.

k_1	$\frac{\|\frac{\partial u}{\partial \mathbf{n}} - W\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|F - F_G\|}{\|F\|}$	$\frac{\|u - u_G\|}{\|u\|}$	cond(A)	# DOF	# DOF per λ_2	C_{QO}
10	2.49e-2	2.32e-2	1.15e-2	4.19e6	432	4.8	1.03
20	2.70e-2	2.36e-2	1.16e-2	9.19e5	400	2.22	1.04
40	1.87e-2	1.82e-2	8.82e-3	4.89e5	400	1.11	1.02
80	8.64e-3	8.83e-3	4.01e-3	2.15e5	400	0.556	1.01
160	5.72e-3	5.65e-3	2.72e-3	2.49e5	360	0.25	1.04

Table 6.11: Scattering by a triangle with refractive index $n = 1.5 + 0.003125\mathrm{i}$ by plane wave with direction $\mathbf{d}^i = (0, -1)$. Relative L^2 errors, condition numbers, and the number of DOF per wavelength in the approximation space.

absorption is reduced. The value of C_{QO} is close to 1 for all k_1 and hence the Galerkin approximation is producing an approximation extremely close to the best approximation.

To conclude this section, we shall compare the relative errors in the far-field for scattering by a triangle with $n = 1.5 + n_I\mathrm{i}$ for $n_I = 0, 1/320, 1/160, 1/80, 1/40, 1/20$ by a plane wave with direction $\mathbf{d}^i = (0, -1)$. These errors are plotted in Figure 6.12. Also here we present the errors in the Kirchhoff approximation for the far-field pattern. We emphasise that the Kirchhoff approximation is an oft-used technique for many scattering problems so it is of interest to observe that we can improve significantly on this approximation with just a small amount of computational effort.

We see in this plot that the far-field relative error is certainly less than 2% for all k_1 when $n_I > 0$ (and in many cases, much less) and appears in most cases to decrease as k_1

150
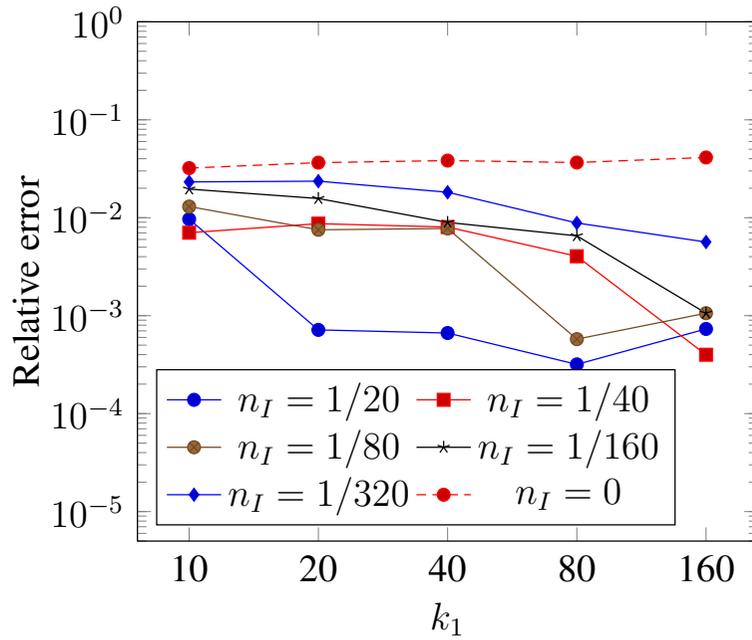
(a) HNA Galerkin BEM approximation with 400 or fewer DOF.
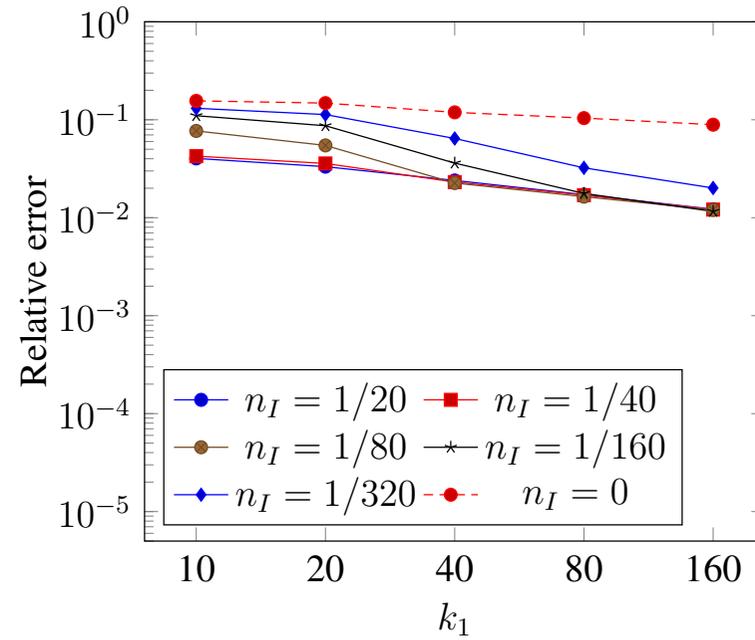
(b) Kirchhoff approximation.

Figure 6.12: Comparison of the HNA Galerkin approximation and the Kirchhoff approximation. Relative L^2 errors in F for scattering of a plane wave with $\mathbf{d}^i = (0, -1)$ by a triangle with refractive index $n = 1.5 + n_I \mathrm{i}$.

increases. This decrease in error as k_1 increases is to be expected for three reasons. Firstly, the GO approximation becomes more accurate, secondly the sizes of discontinuities introduced by the BTA diminish, and thirdly the HNA ansatz becomes more appropriate since it is asymptotic (in k_1) in its nature. For $n_I = 0$, however, we do not see an improvement as k_1 increases. This is due to the algebraic rather than exponential decay of the neglected components as $k_1 \to \infty$. Nevertheless, neither do we see a strong deterioration in the accuracy with an error for this case, with an error of around 3% being maintained for all k_1. So we see that the HNA approach proposed here produces a reasonable (albeit not better than our imposed tolerance of 2%) accuracy for scattering by a triangle with zero absorption.

When comparing to the Kirchhoff approximation, we see that the HNA approximation always provides an improvement but with this improvement being most stark for the higher absorption examples since this is where the approximation space captures the vast majority of the oscillatory behaviour of the diffracted field. Nevertheless, for low absorption and for zero absorption, the HNA method at worst halves the error achieved by the KA with a #DOF of 400 or less.

6.3.2   Scattering by polygons with more than three sides

Up to now we have focused on the problem of scattering by an equilateral triangle. However, the algorithms presented are applicable to general convex polygons. In this section we demonstrate this by presenting results for two examples: scattering by a square with $n = 1.2964 + 0.037\mathrm{i}$ and scattering by a hexagon with $n = 1.3924 + 6.672 \times 10^{-3}\mathrm{i}$. We shall see that the HNA BEM works well for both but that, in the hexagon example (although it could occur for any polygon with $n_s > 3$), the linear system can become ill-conditioned owing to small elements arising in the $\mathrm{e}^{ik_2 r}$ mesh. We shall explain the causes and remedies for this in detail in §6.3.2.2. Firstly we shall consider the square example for which the imaginary part of the refractive index is large enough that this ill-conditioning problem does not arise.

Also, note that we are using refractive indices different to those considered before. This is to illustrate that the algorithms presented are also applicable to scatterers with a range of refractive index. The chosen values $n = 1.2964 + 0.037\mathrm{i}$ and $n = 1.3924 + 6.672 \times 10^{-3}\mathrm{i}$ correspond to the refractive index of ice for light at the wavelengths $8.333\mu m$ and $3.732\mu m$, respectively, as taken from [140]. These wavelengths lie, respectively, within the detecting bandwidths of the IR 8.7 and IR 3.9 sensing channels of SEVIRI, a second generation Meteosat satellite for meteorological observations [45]. Therefore these are refractive indices of interest to meteorologists.

6.3.2.1 Scattering by a square

We consider the scattering of a plane wave travelling at an angle of $3\pi/4$ radians to the x-axis by a square of side length 2π and refractive index $n = 1.2964 + 0.037\text{i}$, as depicted in Figure 6.13. We employ the HNA BEM with $p = 3$ and note that the absorption is large enough such that the $\text{e}^{\text{i}k_2 r}$ basis functions are not present for this example according to the tolerance (5.11) imposed in §5.4. As can be seen in Fig. 6.13 the strong absorption created a weak field within the scatterer and behind it. Nevertheless, diffracted waves can clearly be seen propagating along the sides not illuminated by the incident wave.
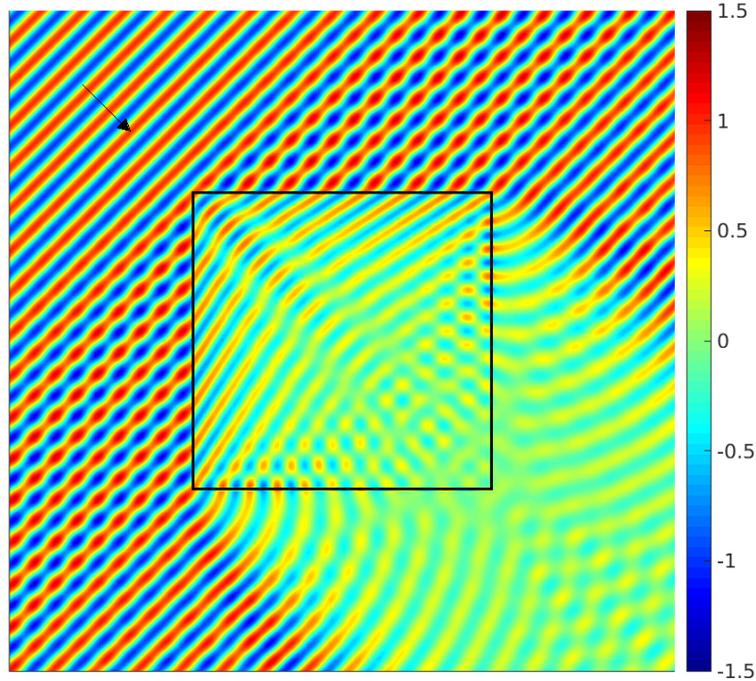
Figure 6.13: Real part of the total field $u = u^i + u^s$ for scattering by a square with $n = 1.2964 + 0.037\text{i}$ by a plane wave with $k_1 = 10$ and incident direction $\mathbf{d}^i = \frac{1}{\sqrt{2}}(1, -1)$.

Table 6.14 presents the relative L^2 errors in the approximation of $\partial u/\partial \mathbf{n}$ on Γ. Those for u on Γ are similar and follow the same trend. We see that the accuracy is better than 2% for all k_1. In fact, the error is substantially better than this for $k_1 \geq 20$. The number of degrees of freedom in the approximation space and the number of degrees of freedom per wavelength (λ_2) are also presented. We see that the number of DOF per λ_2 is below 5 for $k_1 = 10$ and hence comparable to the efficiency of the conventional BEM (see §3.2.4), and for $k_1 = 160$ we achieve an accuracy of less than 0.1% with the number of DOF per

153

λ_2 equal to 0.251 which is a substantial saving compared to a conventional method. We also present the condition numbers of the matrix A and notice that they are well-behaved with a value around 10^6 for all k_1 considered. Finally, we observe that the quasi-optimality

k_1	$\dfrac{\lVert \frac{\partial u}{\partial \mathbf{n}} - W \rVert}{\lVert \frac{\partial u}{\partial \mathbf{n}} \rVert}$	$\dfrac{\lVert \frac{\partial u}{\partial \mathbf{n}} - \left(\frac{\partial u}{\partial \mathbf{n}}\right)_{go} \rVert}{\lVert \frac{\partial u}{\partial \mathbf{n}} \rVert}$	cond(A)	# DOF	# DOF per λ_2	C_{QO}
10	1.70e-2	2.01e-1	1.68e6	480	4.63	1.01
20	2.55e-3	1.44e-1	1.19e6	480	2.31	1.01
40	1.19e-3	1.02e-1	3.52e6	480	1.16	1.04
80	1.28e-3	7.26e-2	5.36e5	416	0.501	1.05
160	8.40e-4	5.14e-2	4.08e5	416	0.251	1.02

Table 6.14: Relative L^2 errors in the approximation of $\partial u/\partial \mathbf{n}$ on Γ by the HNA BEM and GO, condition numbers, degrees of freedom information, and quasi-optimality ratio.

ratio remains close to 1 for all k_1 with little deviation. When we contrast this to the triangle case, it would suggest that the quasi-optimality ratio is geometry dependent, with sharper corners producing a larger ratio.
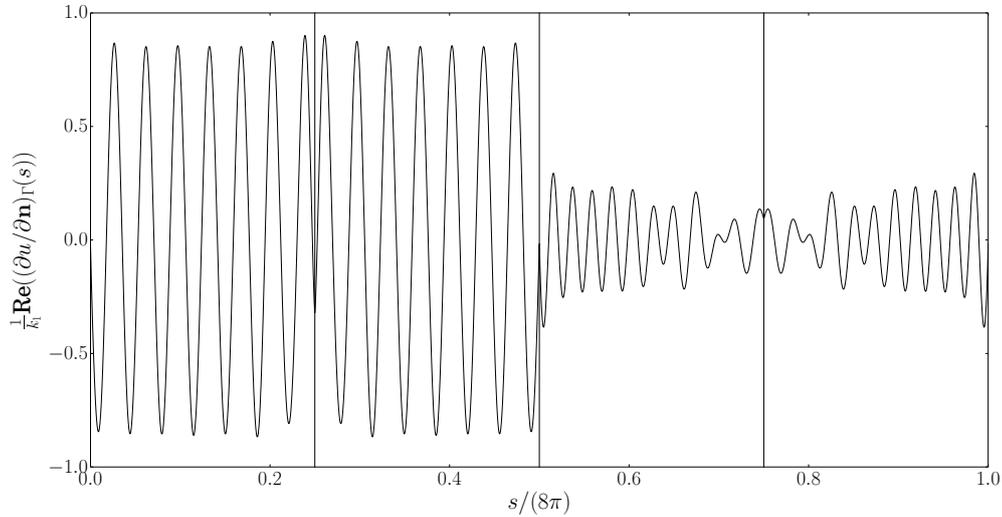
Figure 6.15: Real part of $\partial u/\partial \mathbf{n}$ (scaled by k_1) on Γ for $k_1 = 10$. The vertical lines represent the corners of the square.

In order to understand why the method produces such good approximations for this example, it is helpful to look at the field on the boundary which is plotted for $k_1 = 10$ and $k_1 = 80$ in Figures 6.15 and 6.16, respectively. From the first to the second figure the frequency of course increases by a factor of eight, but also of note is that the field in the shadow region becomes extremely small as k_1 is increased. So small in fact, that diffracted

154
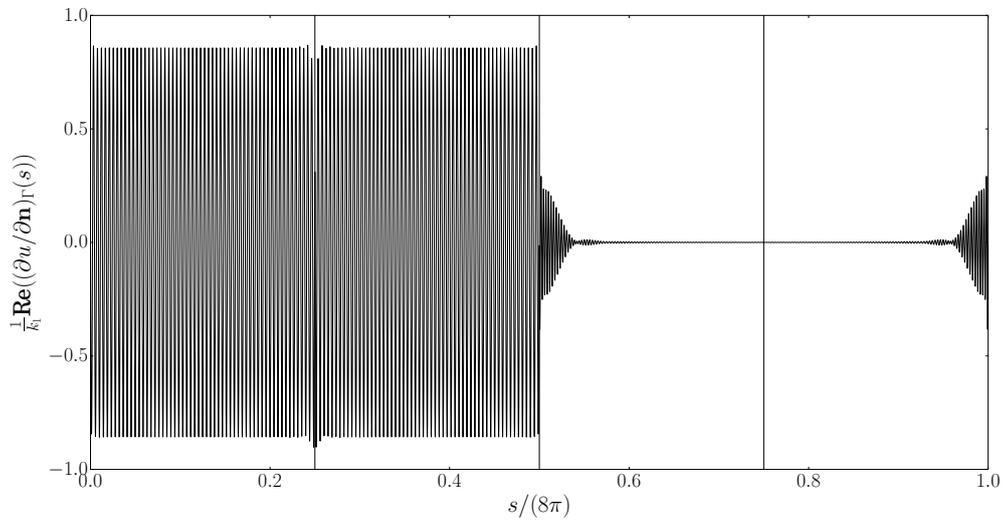
Figure 6.16: Real part of $\partial u/\partial \mathbf{n}$ (scaled by k_1) on Γ for $k_1 = 80$. The vertical lines represent the corners of the square.

waves propagating within the shape decay so rapidly that they do not reach the other sides, let alone reflect. Therefore, we can be confident that, in this case, the proposed HNA ansatz is capturing a large proportion of the diffracted wave behaviour.

6.3.2.2 Scattering by a hexagon

We consider the scattering of a plane wave travelling at an angle of $\tan(2/3)$ radians to the negative x-axis by a hexagon of side length 2π and refractive index $n = 1.3924 + 6.672 \times 10^{-3}\,\mathrm{i}$, as depicted in Figure 6.17. We employ the HNA BEM with $p = 2$. For this example, the absorption is low enough that the $\mathrm{e}^{\mathrm{i}k_1 r}$-type basis functions are present in the approximation space.
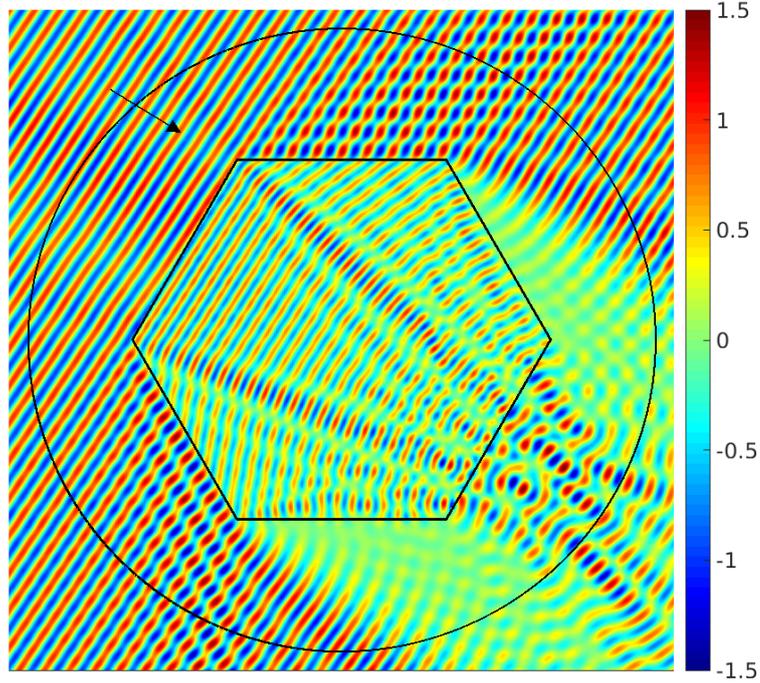
Figure 6.17: Real part of the total field $u = u^i + u^s$ for scattering by a hexagon with $n = 1.3924 + 6.672 \times 10^{-3}\mathrm{i}$ by a plane wave with $k_1 = 10$ and direction $\mathbf{d}^i = (\cos\theta^i, -\sin\theta^i)$ where $\theta^i = \tan^- 1(2/3)$.

We employ the HNA BEM as described up to this point and present the results in Table 6.18. We see that for all k_1 the error in the *best approximation* to $\partial u/\partial \mathbf{n}$ is approximately 2% or less. However, for $k_1 = 10$, the error achieved by the HNA BEM is approximation 28%. This is caused by the extremely poor conditioning of the matrix A, which for this value of k_1 has a condition number of 3.18×10^{18}. The condition number of A for $k_1 = 20$ is also extremely large, however the solution obtained in this case is still close to the best approximation. The cause of this ill-conditioning is the presence of elements in the $\mathrm{e}^{\mathrm{i}k_1 r}$ mesh which are small enough so that the $\mathrm{e}^{\mathrm{i}k_1 r}$-type basis functions are

k_1	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-W_{BA}\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	$\frac{\|\frac{\partial u}{\partial \mathbf{n}}-\left(\frac{\partial u}{\partial \mathbf{n}}\right)_{go}\|}{\|\frac{\partial u}{\partial \mathbf{n}}\|}$	cond(A)	# DOF	# DOF per λ_2	C_{QO}
10	2.77e-1	1.60e-2	2.17e-1	3.17e18	888	5.15	17.3
20	2.19e-2	2.02e-2	1.17e-1	5.99e17	816	2.44	1.08
40	1.18e-2	1.26e-2	6.41e-2	9.75e4	672	1.01	1.02
80	2.28e-3	2.22e-3	4.39e-2	4.69e4	528	0.395	1.03
160	2.95e-3	2.72e-3	3.71e-2	1.33e5	528	0.198	1.10

Table 6.18: Relative L^2 errors in the approximation of $\partial u/\partial \mathbf{n}$ on Γ by the HNA BEM and GO, condition numbers, degrees of freedom information, and quasi-optimality ratios.
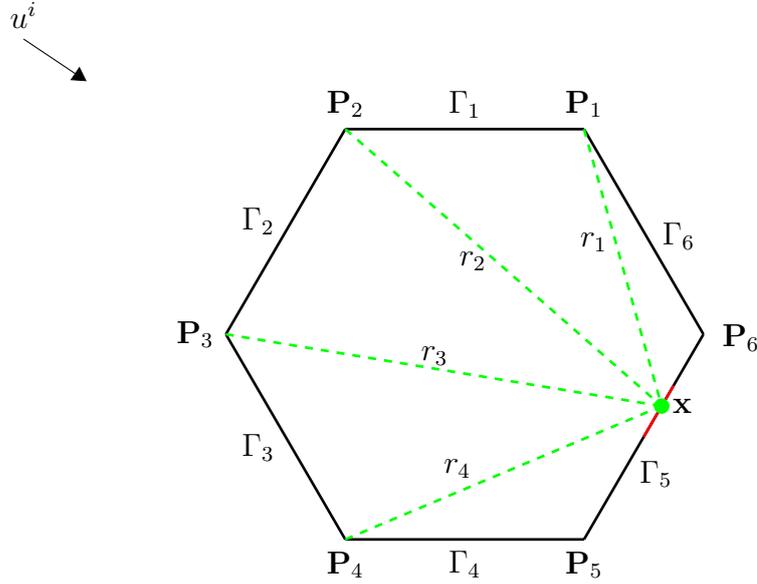
Figure 6.19: The distances r_1, r_2, r_3, r_4 from their respective corners $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3, \mathbf{P}_4$ to $\mathbf{x} \in \Gamma_5$.

not oscillatory across them.

To understand precisely why this can lead to ill-conditioning, let us begin by detailing the approximation space on the side Γ_5, say, in Figure 6.19. On the side Γ_5 we have five meshes as shown in Figure 6.20. Four of these are graded towards the corners and one has elements dictated by strong (in the sense described in §5.3.1 and defined by the tolerance (5.9)) shadow boundaries introduced by the beam tracing algorithm. On each of the elements in this latter mesh, there are 12 basis functions (in general, $(n_s - 2)(p + 1)$ basis functions, where p is the maximum polynomial degree and n_s is the number of sides). These are split into 4 ($n_s - 2$, in general) groups of $p + 1$ functions corresponding to the different oscillatory functions

$$P_i(r_1)\exp(\mathrm{i}k_2r_1),\ P_i(r_2)\exp(\mathrm{i}k_2r_2),\ P_i(r_3)\exp(\mathrm{i}k_2r_3),\ P_i(r_4)\exp(\mathrm{i}k_2r_4) \quad i = 0, \ldots, n_s{-}3,$$
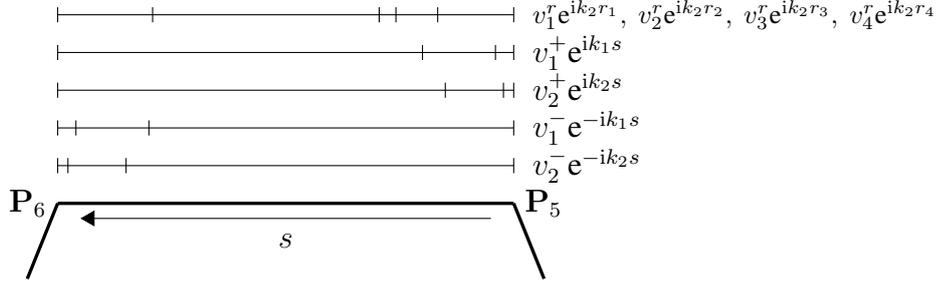
Figure 6.20: Illustration of overlapping meshes used to approximate the amplitudes v_1^+, v_2^+, v_1^-, v_2^-, v^r associated with the phase functions on a typical side Γ_5.

where the distances r_1, r_2, r_3, r_4 are distances from corners $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$, $\mathbf{P}_4$, respectively, as shown in Figure 6.19. When k_1 is large relative to the size of a particular element, these functions will oscillate many times across the length of the element and hence the different phases will enable the functions to correctly capture different oscillatory solution components. However, when k_1 is small relative to the size of the element, these functions closely resemble one another, so that the approximation space has redundant basis functions. This is the cause of the ill-conditioning (and unacceptable approximation error) for small k_1 values in Table 6.18.

In order to remedy this situation we may introduce a rule by which we discard basis functions if they are not sufficiently oscillatory on an element. In order to formulate such a rule, we parameterise the distance functions r_i, $i = 1, \ldots, n_s - 2$ by the arc length s along the side Γ_5. Therefore $r_i(s)$ can be written as

$$r_i(s) = \sqrt{s^2 + 2As + B}, \tag{6.15}$$

where $A = \mathbf{t}_5 \cdot (\mathbf{P}_5 - \mathbf{P}_i)$ and $B = |\mathbf{P}_5 - \mathbf{P}_i|^2$. We note that this function has a stationary point when $s = -A$. Consider an element which is marked in red on Γ_5 in Figure 6.19 and is the interval $s = [a, b]$. Then we may say that the function $\exp(ik_2 r_i)$ is non-oscillatory on $[a, b]$ if either

- $-A \notin [a, b]$ (i.e., not stationary point on the interval) and $|r_i(a) - r_b(a)| < \epsilon \frac{2\pi}{\text{Re}(k_2)}$,

or

- $-A \in [a, b]$ and $|r_i(a) - r_i(-A)| + |r_i(-A) - r_i(b)| < \epsilon \frac{2\pi}{\text{Re}(k_2)}$,

where ϵ is a tuning parameter to be chosen. Simply stated, this says that we define r_i as non-oscillatory if it completes fewer than ϵ oscillations across $[a, b]$.

If a basis function is classified as not oscillatory according to the above definition then we discard it. However, if all the basis functions on an element are discarded by this

process, then we simply reintroduce the first $p+1$ Legendre polynomials as (conventional) basis functions on this element. In this way, we can eliminate redundant basis functions which give rise to ill-conditioning.

Numerical experiments suggest that for the case considered here, the parameter choice $\epsilon = 3/2$ provides a good compromise between conditioning and accuracy. A smaller value of ϵ leads to better accuracy but worse conditioning, and a larger value leads to poorer accuracy but better conditioning. The results for the method after this additional piece of the algorithm is included are displayed in Table 6.22. Along with the updated results of Table 6.18, also presented are relative errors in the total field on the circle (radius 3π) in Figure 6.17 and in the far-field. We also provide these errors for the GO approximation for comparison.

The table shows that now, for all k_1, we achieve an accuracy of roughly 2% or better, and the HNA BEM produces a solution close to the best approximation. The far-field accuracy is slightly better than that on the boundary and the accuracy in the total field is better still, with values less than 1% for all k_1 considered. The relationship between these errors is clearly seen in the plot Figure 6.21. The new values of C_{QO} are much closer to 1 than previously and the condition numbers are much smaller for the problematic $k_1 = 10$ and $k_1 = 20$ examples.
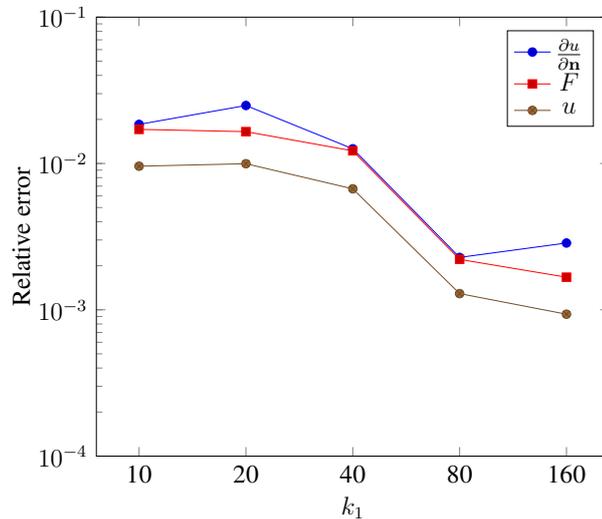
Figure 6.21: Relative L^2 errors in $\partial u_\Gamma / \partial \mathbf{n}$, F_G, and u_G on the circle in Figure 6.17 for hexagon with $n = 1.3924 + 6.72 \times 10^{-3}$i.

When comparing to the GO approximation on the boundary and the Kirchhoff approximation in the domain Ω_1 and the far-field, we observe a significant improvement across the frequency range considered using just 816 or few degrees of freedom. For example, for

159

$k_1 = 10$, the error in the far-field is 1.71% for the HNA BEM and 19.3% for the Kirchhoff approximation, and for $k_1 = 160$, the error in the far-field is 0.167% for the HNA BEM and 1.85% for the Kirchhoff approximation.

To conclude, we provide plots of the far-field pattern in Figure 6.23 and Figure 6.24 for $k_1 = 20$ and $k_1 = 160$. The increased oscillatory nature of the field as well as the strengthening of the shadow and reflected regions as k_1 increases are clearly noticeable.

Figure 6.25 and Figure 6.26 show the total field on the circle in Figure 6.17. Again, we observe that for $k_1 = 20$, the field is of a similar magnitude all around the circle, indicating that diffraction plays an important role in this case. For $k_1 = 160$, the shadow and reflected regions, which are GO phenomena, are very pronounced. This later observation implies that in this case, the GO approximation should give a good approximation which, as can be seen in Table 6.22, is indeed the case. Nevertheless, the HNA BEM provides an even more accurate approximation with less than 0.2 DOF per wavelength for this example.

k_1	Approximations on Γ						Far-field		In Ω_1					
	$\frac{\lVert u-U\rVert}{\lVert u\rVert}$	$\frac{\lVert\frac{\partial u}{\partial \mathbf{n}}-W\rVert}{\lVert\frac{\partial u}{\partial \mathbf{n}}\rVert}$	$\frac{\lVert u-U_{BA}\rVert}{\lVert u\rVert}$	$\frac{\lVert\frac{\partial u}{\partial \mathbf{n}}-W_{BA}\rVert}{\lVert\frac{\partial u}{\partial \mathbf{n}}\rVert}$	$\frac{\lVert u-u_{go}\rVert}{\lVert u\rVert}$	$\frac{\lVert\frac{\partial u}{\partial \mathbf{n}}-\left(\frac{\partial u}{\partial \mathbf{n}}\right)_{go}\rVert}{\lVert\frac{\partial u}{\partial \mathbf{n}}\rVert}$	$\frac{\lVert F-F_G\rVert}{\lVert F\rVert}$	$\frac{\lVert F-F_{go}\rVert}{\lVert F\rVert}$	$\frac{\lVert u-u_2\rVert}{\lVert u\rVert}$	$\frac{\lVert u-u_{go}\rVert}{\lVert u\rVert}$	cond(A)	# DOF	# DOF per λ_2	C_{QO}
10	2.53e-2	1.85e-2	2.01e-2	1.62e-2	3.55e-1	2.17e-1	1.71e-2	1.93e-1	9.58e-3	1.12e-1	8.91e8	762	4.56	1.14
20	4.40e-2	2.49e-2	3.91e-2	2.36e-2	2.07e-1	1.17e-1	1.65e-2	8.21e-2	9.96e-3	5.12e-2	1.43e7	816	2.44	1.06
40	2.46e-2	1.26e-2	2.26e-2	1.18e-2	1.16e-1	6.41e-2	1.22e-2	4.17e-2	6.71e-3	2.40e-2	9.75e4	672	1.01	1.07
80	4.67e-3	2.28e-3	4.52e-3	2.22e-3	7.82e-2	4.39e-2	2.21e-3	2.62e-2	1.29e-3	1.52e-2	4.69e4	528	0.395	1.03
160	4.14e-3	2.86e-3	3.50e-3	2.73e-3	5.55e-2	3.11e-2	1.67e-3	1.85e-2	9.33e-4	1.07e-2	1.33e5	528	0.198	1.06

Table 6.22: Relative L^2 errors for scattering by a hexagon. Also presented are the condition number of the matrix A, the number of degrees of freedom in the approximation space, and the number of degrees of freedom per wavelength around the boundary. Here the redundant basis functions have been removed (with $\epsilon = 3/2$). Compare to before this removal in Table 6.18.
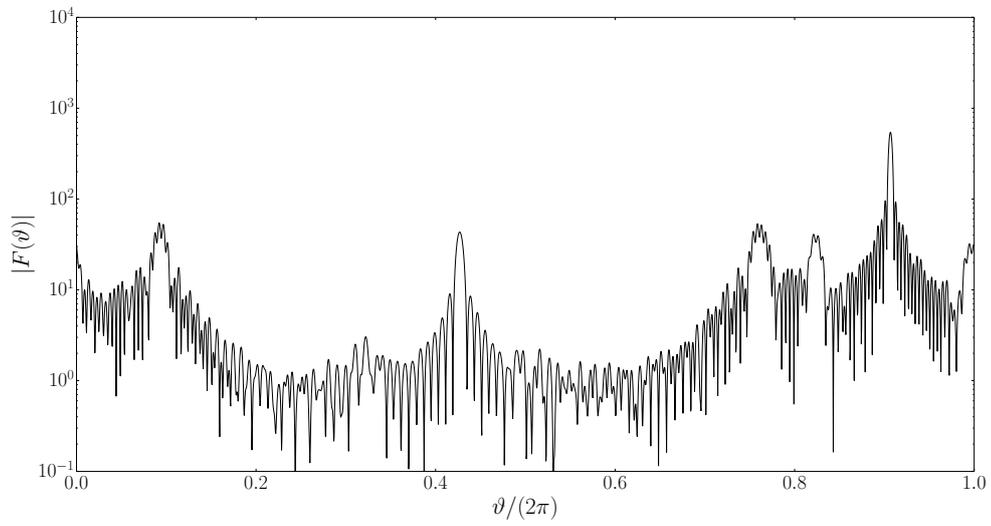
Figure 6.23: Far-field pattern for $k_1 = 20$.
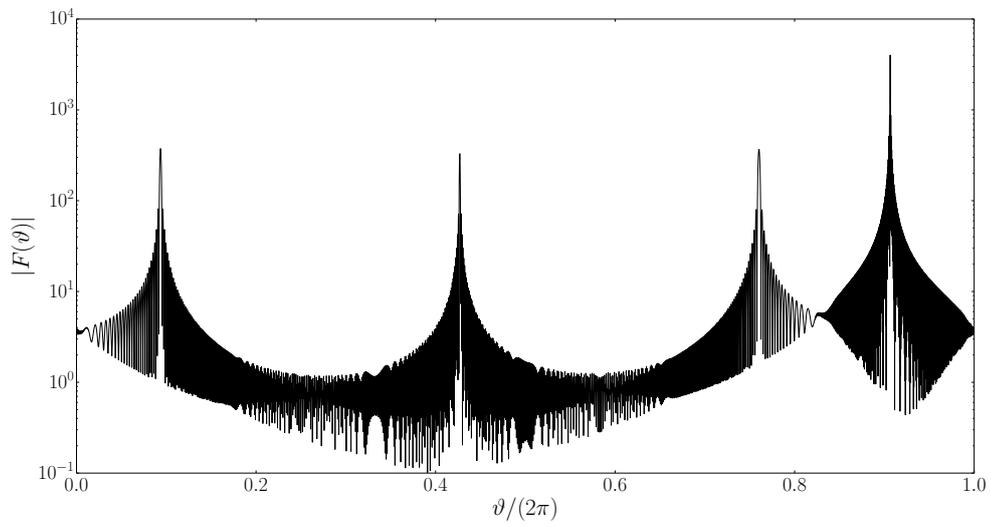
Figure 6.24: Far-field pattern for $k_1 = 160$.

162
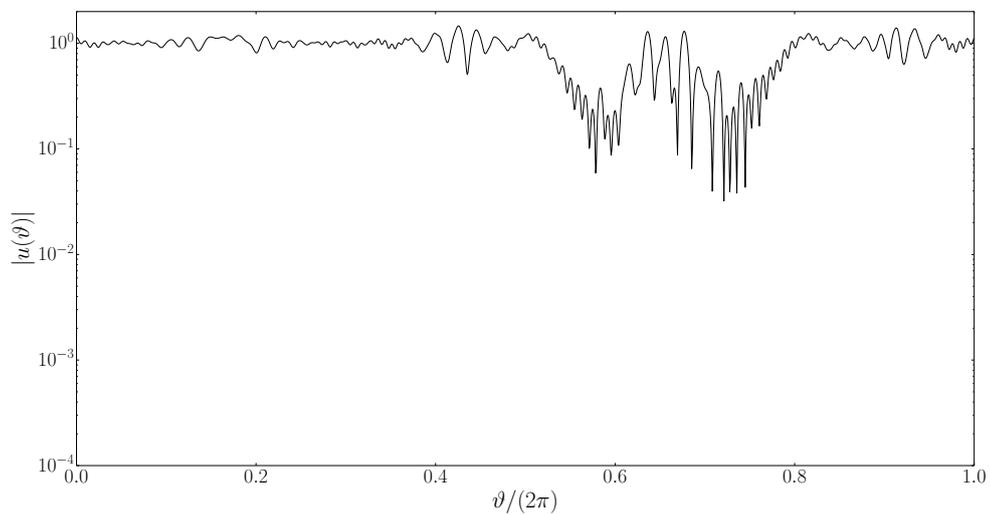
Figure 6.25: Total field on the circle in Figure 6.17 pattern for $k_1 = 20$.
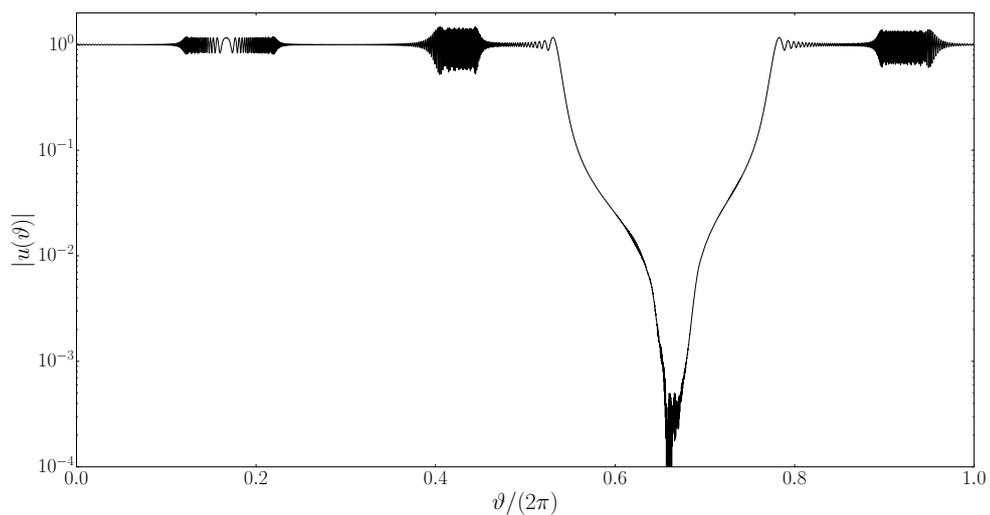
Figure 6.26: Total field on the circle in Figure 6.17 pattern for $k_1 = 160$.

Chapter 7

Conclusions

The main aim of this thesis has been to extend the hybrid numerical-asymptotic (HNA) approach for the first time to scattering by penetrable obstacles. This task has required us to address numerous asymptotic and numerical aspects associated with the solution of such scattering problems. In this chapter, we summarise the main contributions of this thesis to tackle these aspects and develop an HNA BEM for penetrable obstacles. We conclude the chapter by indicating areas of potential future research for the improvement and extension of the method developed here, as well as HNA methods as a whole.

7.1   Summary of results

In §2 we began by stating the 3D EM transmission scattering problem to be solved along with its reformulation as a system of boundary integral equations. We showed how in the 2D case, this EM problem boils down to solving two 2D acoustic transmission problems, one for each of the electric and magnetic fields. We then presented the reformulation of the 2D acoustic transmission problem as a system of boundary integral equations. We went on to prove, using a compact perturbation argument, that the integral equations are uniquely solvable.

Next in §3 we considered the numerical solution of the BIEs given in §2. We presented a brief comparison of the performance of a state-of-the-art implementation of a conventional BEM with a T-matrix method for the solution of the 3D EM scattering problem. We saw that the memory and time requirements of both methods scale extremely poorly with increasing size parameter and that only problems of size parameter less than 20 can be solved on a standard desktop (if an accuracy of approximation 1% in the far-field is required). This demonstrated the limitations of current standard techniques and hence provided motivation for the development of the HNA method of this thesis. §3.2 onwards discussed the implementation of a 2D conventional Galerkin BEM for the transmission problem. It was seen

that approximately six degrees of freedom per wavelength λ_1 (or four per λ_2) were required to achieve approximately 1% accuracy in the far-field and hence for high-frequency problems, many thousands of degrees of freedom are required to accurately approximate the solution.

The first important aspect of our HNA method was the development of a beam tracing algorithm for the calculation of the geometrical optics approximation. This was the topic of §4.1–§4.6. In order to develop the BTA, we considered the canonical problem of the scattering of a plane wave at a planar interface between two absorbing media. The same problem but for two non-absorbing media is completely classical and its analysis appears in standard physics textbooks. The absorbing media case, however, has been far less well studied and in fact has led to some contradictory analyses in the literature. Therefore, we derived the formulae afresh in a way well-suited to the implementation of a BTA, that is, in terms of vectors rather than angles relative to the interface. We found that the formulae for the direction of the transmitted direction vector contains a square root whose sign must be specified. In order to decide which sign choice to make, we compared the GO approximation obtained from a simple scattering problem to a reference solution for different incident angles. We found that the correct sign choice depends on how close the incident wave is to being totally internally reflected and that there appear to be two regimes, each with its own sign choice. The two regimes are connected by a small (and shrinking as k increases) region in which neither sign choice gives a good approximation and hence our GO approximation appears to be invalid. The observation of this phenomenon appears to be novel. We developed a rule of thumb for switching between the two choices resulting in a consistent (in terms of incident wave direction) BTA for scattering by convex penetrable polygons.

In §5 we developed the foundation of our HNA BEM, namely the HNA approximation space for the transmission problem. Our specific aim in this thesis was to achieve a prescribed relative error tolerance in the boundary solution of 2%. We stress that a smaller or larger error tolerance could have been chosen, but this would have resulted in a larger approximation space as well as different (tuning) parameter values in the method. We showed via least squares fitting to a reference solution that for scattering by equilateral triangles in which $k_1 n_I \geq 0.125$ (where n_I is the imaginary part of the refractive index), and for the two incident wave directions considered there, our 2% error tolerance could be achieved with a small number of degrees of freedom which does not grow as k_1 increases. For problems where $k_1 n_I < 0.125$ the method does not achieve this error tolerance for all k_1 considered, nevertheless it achieves approximations which are significantly more accurate than GO (which we emphasise is a commonly employed approximation) using only a small

number of degrees of freedom. It was seen for experiments with a square and hexagon that indeed the approximation space is suitable for general convex polygons and also that for these shapes, since they have a larger interior than the triangle, the approximation space performs even better for lower absorptions than it did for the triangle. Finally, we looked at the convergence of the method as the polynomial degree is increased.

The HNA approximation space developed in §5 was implemented within a Galerkin boundary element method in §6. In order to perform this implementation accurately, one must take care to evaluate the oscillatory integrals accurately, as well as normalise the basis functions appropriately. Results were presented demonstrating the method's efficacy for scattering by a triangle, square and hexagon with various refractive indices. We saw that close to the best approximation results of §5 were achieved for the scattering by a triangle examples. In addition to considering some of the same scattering setups from §5, we considered an example where the incident wave grazes the side Γ_1. For this example it was seen that the 2% tolerance was achieved for $k_1 n_I \geq 0.5$ rather than for $k_1 n_I \geq 0.125$ as for the other two incident directions. We hypothesised that this was due to a prominent head wave generated by the incident wave at grazing. This hypothesis was supported by the regular oscillatory nature of the error in the approximation on Γ_3 where we expect to see the influence of the head wave. This suggests that the current HNA BEM is not completely robust to incident wave direction and that at grazing incidence, one may have to include basis functions with the phase of the anticipated head wave. Further investigation is required to ascertain for certain whether indeed this larger error is in fact due to the head wave however.

Further to the equilateral triangle examples, scattering by a square and a regular hexagon were considered. We chose refractive indices relevant to the application to light scattering by atmospheric ice crystals. We saw that, for the hexagon, the method achieves better than 1% accuracy in the domain for all k_1 considered with a number of degrees of freedom of 816 or fewer. For this example with $k_1 = 160$, we achieved an error in the domain of 0.09% with 0.2 degrees of freedom per wavelength. When compared to the common-place requirement of 6 to 10 degrees of freedom per wavelength for "engineering accuracy" (a few percent) with conventional methods, we can see that the HNA BEM allows one to significantly reduce the number of degrees of freedom required to solve transmissions problems. Further, we observed in all the examples we considered that our HNA BEM offers an improvement in accuracy of approximation: a factor of 10 over the commonly-employed Kirchhoff approximation.

7.2 Future work

This section is divided into two components, namely the future work relevant to only the HNA BEM for the transmission problem, and the future work applicable to the HNA approach as a whole. We shall begin by discussing the former.

7.2.1 The HNA approach for the transmission problem

The HNA BEM developed in this thesis is the first HNA method for scattering by a penetrable scatterer, with all other HNA BEMs in the literature being applicable only to the impenetrable case. As we have seen, this simple change in the definition of the scattering problem leads to a huge increase in the complexity of the high-frequency behaviour of the scattered wave field, and hence in the complexity of the associated HNA BEM. We saw that the GO approximation now consists of infinitely many terms in comparison to the single term in the impenetrable convex case. Further, the diffracted field also in principle consists of infinitely many terms compared to the two per side in the impenetrable convex case. In an attempt to simplify the problem, we considered scatterers possessing some absorption. This did indeed simplify the diffracted field however introduced different difficulties concerning the propagation of inhomogeneous waves within the scatterer. We also saw that many of the diffracted terms contain discontinuities introduced by the beam tracing procedure which must be taken into account when constructing the corresponding meshes.

Each of these topics was tackled to an extent in the construction of the HNA method of this thesis. However, for each, there are improvements that could be made which we discuss in turn.

- **Geometrical optics for absorbing media.** When calculating the GO approximation for scattering by absorbing obstacles, we encountered the phenomenon that the occurring inhomogeneous waves are defined by an equation which has two solutions. We saw that, depending on problem parameters (incident wave direction, wavenumber etc.), the seemingly correct solution - determined via comparison to a reference solution - differed. In fact, there appears to be two regions, one for either solution, with a connecting region in which the GO approximation is seemingly invalid. These regions are related to the phenomenon of total internal reflection. This observation appears to be novel and the reasons for its occurrence have not been determined in this thesis. Further research into this phenomenon and a technique to patch the two regions together would improve the accuracy of the GO approximation and hence the overall accuracy of the HNA BEM.

- **Overlapping meshes for the e^{ik_2r} basis functions.** It was chosen here to accommodate all the basis functions of the form $e^{ik_2r_1}, e^{ik_2r_2}, \ldots, e^{ik_2r_{n_s-2}}$ on a single mesh. As we saw in §6.3.2.2, sometimes this single mesh implementation leads to redundancy and hence ill-conditioning if there are small elements present and we must design an algorithm by which to discard basis functions in this case. An overlapping sequence of meshes, each accommodating one of the $e^{ik_2r_1}, e^{ik_2r_2}, \ldots, e^{ik_2r_{n_s-2}}$ type functions would avoid this problem and arguably is a more natural way to design the meshes. Only a small modification to the current implementation would be required to change to the overlapping mesh setup just described. Hence an investigation into the efficacy of the method with this new mesh setup would be simple to undertake in future.

- **Shadow boundary mesh refinement.** To compensate for each discontinuity introduced by the BTA, we added a mesh division in the appropriate numerical mesh. However this limited the accuracy and convergence (in p) of the method. To improve the accuracy and convergence, meshes graded towards these discontinuities, as discussed in [61], should be employed. For a problem with sufficiently high absorption, i.e., one in which only the shadow boundaries from first order reflections are significantly large, we would expect such an approach to yield an hp-HNA BEM which is exponentially convergent in the polynomial degree p, similar to some previous HNA methods (e.g., [63]) .

- **Higher order terms.** In this thesis, we considered only the first order diffracted waves (i.e., not reflections of diffracted waves), and neglected head waves. A natural extension of the method proposed here, which may allow the level of absorption to be lowered without degrading the accuracy, would be to incorporate the reflections of diffracted waves and head waves into the approximation space. Based on the results presented in this thesis, it is likely that the inclusion of basis functions corresponding to one extra reflection of the diffracted waves, as well as the first order heads waves, would lead to a method which achieves the tolerance set in this thesis for scatterers at any absorption. Further, we expect that the inclusion of basis functions corresponding to head waves will make the method more robust to incident wave direction since at grazing incidence, the head waves are most prominent and the current method which neglects them performs less well.

- **Non-convex geometries.** The HNA approximation space proposed here could be extended to non-convex scatterers by the introduction of basis functions which oscillate as $\exp(ik_1r)$ where r is the radial distance from the corresponding corner. These

functions would propagate in the exterior medium (Ω_1) and are the same as those proposed for the HNA method for impenetrable non-convex polygons in [29]. We may also require to consider the transmission of these diffracted waves into the polygon and then reflections within the scatterer. Indeed, the complexity of such problems can quickly escalate to such a level when the HNA method as proposed here needs considerable revision. It is recommended that an investigation into this problem begin with highly absorbing scatterers with geometries such as those in [29] and then try to incorporate further phases when appropriate as the absorption is reduced.

The final point we mention is not necessarily one requiring further research but rather an implementation challenge which will improve the speed (in terms of CPU time) of the transmission HNA BEM.

- **Oscillatory quadrature.** The implementation of the HNA BEM presented in this thesis relied on a composite Gaussian quadrature rule to approximate the arising oscillatory integrals. Therefore, the time required to calculate the mass matrix scaled as k_1^2. Oscillatory quadrature techniques exist which allow the frequency-independent approximation of such 2D integrals. Their implementation is challenging but possible and would lead to an HNA BEM which not only has a number of degrees of freedom independent of frequency, but also a computation time which is short and independent of frequency.

7.2.2 The general HNA methodology

A thorough review of the HNA approach is given in [28] with many of its current challenges and future directions discussed in detail. Some of these challenges pertain to the theoretical aspects of the numerical approximation of the scattering problems, e.g., k-explicit coercivity estimates for integral operators. Here, however, we shall focus on the forward directions with regard to the development and implementation of practical and efficient HNA BEMs.

- **Collocation HNA BEM.** As was mentioned in §6.2.2 the efficient numerical evaluation of the oscillatory integrals arising in the HNA method is a significant challenge. Many methods exist and are well-known however their implementation is not a trivial task for the 2D integrals arising in the Galerkin method for 2D problems. With the progression to 3D problems, the Galerkin implementation would require the evaluation of 4D oscillatory integrals, the implementation of which may prove to not be feasible or at least practical. Therefore, the development of efficient collocation versions of present HNA methods is an attractive avenue for future research. We point

out that collocation HNA h-BEMs have been presented for some scattering problems (e.g., [4]), however research in collocation HNA hp-BEMs is still in it early stages with the study in [115] providing promising results for scattering by screens in 2D.

- **Multiple scattering configurations.** HNA methods have been considered for multiple scattering problems involving smooth scatterers (see, e.g., [43]) for which the techniques proposed rely on an iterative approach which decomposes the problem into many single scattering problems. This approach, to the author's knowledge, has not been extended to the polygonal scatterer case and hence is a natural extension for future research. However we note that a different HNA method for multiple scatterers has been proposed that does not rely on iteration [27].

- **Three-dimensional problems.** Finally, the ultimate aim of the HNA methodology would be to efficiently solve problems of relevance in science and engineering, namely problems in three dimensions. This extension is a challenging task since, for the transmission problem, each 2D face of the polyhedron corresponds to a form of the 2D problem discussed in this thesis, and we also have to take account of the 3D diffracted waves propagating within the scatterer. Some work has been done for scattering by a 2D screen in three dimensions [58] which is a natural starting point for 3D problems. The results are promising and with further research could lead the way for HNA methods for scattering by 3D polyhedra.

Bibliography

[1] R. A. ADAMS AND J. J. F. FOURNIER, *Sobolev Spaces*, Academic Press, 2003.

[2] Y. A. ANTIPOV AND V. V. SILVESTROV, *Diffraction of a plane wave by a right-angled penetrable wedge*, Radio Sci., 42 (2007), p. RS4006.

[3] X. ANTOINE AND H. BARUCQ, *Approximation by generalised impedance boundary conditions of a transmission problem in acoustic scattering*, ESAIM Math. Model. Numer. Anal., 39 (2005), pp. 1041–1059.

[4] S. ARDEN, S. N. CHANDLER-WILDE, AND S. LANGDON, *A collocation method for high-frequency scattering by convex polygons*, J. Comput. Appl. Math., 204 (2007), pp. 334–343.

[5] K. E. ATKINSON, *The Numerical Solution of Integral Equations of the Second Kind*, Cambridge University Press, 1997.

[6] V. M. BABICH, M. A. LYALINOV, AND V. E. GRIKUROV, *Diffraction theory: the Sommerfeld-Malyuzhinets technique*, Alpha Science International Limited, 2007.

[7] A. J. BARAN, *A review of the light scattering properties of cirrus*, J. Quant. Spectrosc. Ra., 110 (2009), pp. 1239–1260.

[8] ——, *From the single-scattering properties of ice crystals to climate prediction: A way forward*, Atmos. Res., 112 (2012), pp. 45–69.

[9] P. E. BARBONE AND O. MICHAEL, *Scattering from submerged objects by a hybrid asymptotic-boundary integral equation method*, Wave motion, 29 (1999), pp. 137–156.

[10] A. H. BARNETT AND T. BETCKE, *Stability and convergence of the method of fundamental solutions for helmholtz problems on analytic domains*, J. Comput. Phys., 227 (2008), pp. 7003–7026.

[11] M. BEBENDORF, *Hierarchical Matrices: A Means to Efficiently Solve Elliptic Boundary Value Problems*, Springer, 2008.

[12] H. L. BERTONI, A. C. GREEN, AND L. B. FELSEN, *Shadowing an inhomogeneous plane wave by an edge*, J. Opt. Soc. Amer., 68 (1978), pp. 983–988.

[13] T. BETCKE, S. ARRIDGE, J. PHILLIPS, M. SCHWEIGER, AND W. ŚMIGAJ, *Solution of electromagnetic problems using BEM++*, in Oberwolfach Report 03/2013, 2013, pp. 146–150.

[14] T. BETCKE AND E. A. SPENCE, *Numerical estimation of coercivity constants for boundary integral operators in acoustic scattering*, SIAM Journal on Numerical Analysis, 49 (2011), pp. 1572–1601.

[15] L. BI AND P. YANG, *Physical-geometric optics hybrid methods for computing the scattering and absorption properties of ice crystals and dust aerosols*, in Light Scattering Reviews 8, Springer, 2013, pp. 69–114.

[16] L. BI, P. YANG, G. W. KATTAWAR, Y. HU, AND B. BAUM, *Scattering and absorption of light by ice particles: Solution by a new physical-geometric optics hybrid method*, J. Quant. Spectrosc. Ra., 112 (2011), pp. 1492–1508.

[17] J. BILLINGHAM AND A. C. KING, *Wave Motion*, Cambridge University Press, 2000.

[18] C. F. BOHREN AND D. R. HUFFMAN, *Absorption and Scattering of Light by Small Particles*, John Wiley & Sons, Inc, 1983.

[19] M. BORN AND E. WOLF, *Principles of Optics*, Cambridge University Press, 1997.

[20] V. A. BOROVIKOV AND B. Y. KINBER, *Geometrical Theory of Diffraction*, IEE, 1994.

[21] A. G. BOROVOI AND I. A. GRISHIN, *Scattering matrices for large ice crystal particles*, J. Opt. Soc. Am. A, 20 (2003), pp. 2071–2080.

[22] L. M. BREKHOVSKIKH, *Waves in Layered Media*, Academic Press, 1960.

[23] O. BRUNO, C. GEUZAINE, J. MONRO JR, AND F. REITICH, *Prescribed error tolerances within fixed computational times for scattering problems of arbitrarily high frequency: the convex case*, Philos. T. Roy. Soc. A, 362 (2004), pp. 629–645.

172

[24] B. BUDAEV AND D. BOGY, *Rigorous solutions of acoustic wave diffraction by penetrable wedges*, J. Acoust. Soc. Am., 105 (1999), p. 74.

[25] A. BUFFA, M. COSTABEL, AND D. SHEEN, *On traces for h (curl, ?) in lipschitz domains*, Journal of Mathematical Analysis and Applications, 276 (2002), pp. 845–867.

[26] A. BUFFA AND R. HIPTMAIR, *Galerkin boundary element methods for electromagnetic scattering*, in Computational Methods in Wave Propagation, M. Ainsworth, ed., Springer, New York, 2003.

[27] S. N. CHANDLER-WILDE, A. J. GIBBS, S. LANGDON, AND A. MOIOLA, *Hybrid numerical asymptotic approximation for multiple scattering problems*, in The 12th International Conference on Mathematical and Numerical Aspects of Wave Propagation, 2015.

[28] S. N. CHANDLER-WILDE, I. G. GRAHAM, S. LANGDON, AND E. A. SPENCE, *Numerical-asymptotic boundary integral methods in high-frequency acoustic scattering*, Acta. Numer., (2012), pp. 89–305.

[29] S. N. CHANDLER-WILDE, D. P. HEWETT, S. LANGDON, AND A. TWIGGER, *A high frequency boundary element method for scattering by a class of nonconvex obstacles*, Numer. Math., 129 (2015), pp. 647–689.

[30] S. N. CHANDLER-WILDE AND S. LANGDON, *A Galerkin boundary element method for high frequency scattering by convex polygons*, SIAM J. Numer. Anal., 45 (2007), pp. 610–640.

[31] S. N. CHANDLER-WILDE, S. LANGDON, AND M. MOKGOLELE, *A high frequency boundary element method for scattering by convex polygons with impedance boundary conditions*, Commun. Comput. Phys., 11 (2012), pp. 575–593.

[32] P. C. Y. CHANG, J. G. WALKER, AND K. I. HOPCRAFT, *Ray tracing in absorbing media*, J. Quant. Spectrosc. Ra., 96 (2005), pp. 327–341.

[33] S. J. CHAPMAN, J. M. H. LAWRY, J. R. OCKENDON, AND R. H. TEW, *On the theory of complex rays*, SIAM review, (1999), pp. 417–509.

[34] P. C. CLEMMOW, *The Plane Wave Spectrum Representation of Electromagnetic Fields*, Pergamon Press Ltd., 1966.

[35] D. COLTON AND R. KRESS, *Integral Equation Methods in Scattering Theory*, Wiley, 1983.

[36] M. COSTABEL AND F. LOUËR, *On the Kleinman-Martin integral equation method for electromagnetic scattering by a dielectric body*, SIAM J. Appl. Math., 71 (2011), pp. 635–656.

[37] M. COSTABEL AND E. STEPHAN, *A direct boundary integral equation method for transmission problems*, J. Math. Anal. Appl., 106 (1985), pp. 367–413.

[38] J. A. DESANTO, *Scalar Wave Theory*, Springer, 1992.

[39] *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, Release 1.0.5 of 2012-10-01. Online companion to [114].

[40] V. DOMINGUEZ, I. G. GRAHAM, AND V. P. SMYSHLYAEV, *A hybrid numerical-asymptotic boundary integral method for high-frequency acoustic scattering*, Numer. Math., 106 (2007), pp. 471–510.

[41] M. G. DUFFY, *Quadrature over a pyramid or cube of integrands with a singularity at a vertex*, SIAM J. Numer. Anal., 19 (1982), pp. 1260–1262.

[42] M. A. DUPERTUIS, M. PROCTOR, AND B. ACKLIN, *Generalization of complex Snell-Descartes and Fresnel laws*, J. Opt. Soc. Amer. A, 11 (1994), pp. 1159–1166.

[43] F. ECEVIT AND F. REITICH, *Analysis of multiple scattering iterations for high-frequency scattering problems. i: The two-dimensional case*, Numer. Math., 114 (2009), pp. 271–354.

[44] B. ENGQUIST AND L. YING, *Fast directional multilevel algorithms for oscillatory kernels*, SIAM Journal on Scientific Computing, 29 (2007), pp. 1710–1737.

[45] EUROPEAN SPACE AGENCY, *Meteosat Second Generation*. http://www.esa.int/esapub/br/br153/br153.pdf, 1999.

[46] E. FABES, M. SAND, AND J. K. SEO, *The spectral radius of the classical layer potentials on convex domains*, in Partial Differential Equations with Minimal Smoothness and Applications, Springer, 1992, pp. 129–137.

[47] J. D. FENTON, *Numerical methods for nonlinear waves*, Adv. Coast. Ocean Eng., 5 (1999), pp. 241–324.

[48] P. R. FIELD, J. HEYMSFIELD, A. BANSEMER, AND C. H. TWOHY, *Determination of the combined ventilation factor and capacitance for ice crystal aggregates from airborne observations in a tropical anvil cloud*, J. Atmos. Sci., 65 (2008), pp. 376–391.

[49] G. GENNARELLI AND G. RICCIO, *A uniform asymptotic solution for the diffraction by a right-angled dielectric wedge*, IEEE T. Antenn. Propag., 59 (2011), pp. 898–903.

[50] E. GILADI, *An asymptotically derived boundary element method for the Helmholtz equation*, J. Comput. Appl. Math., 198 (2007), pp. 52–74.

[51] E. GILADI AND J. B. KELLER, *A hybrid numerical asymptotic method for scattering problems*, J. Comput. Phys., 174 (2001), pp. 226–247.

[52] S. P. GROTH, *Light scattering by penetrable convex polygons*, Master's thesis, University of Reading, 2011.

[53] S. P. GROTH, A. J. BARAN, T. BETCKE, S. HAVEMANN, AND W. ŚMIGAJ, *The boundary element method for light scattering by ice crystals and its implementation in BEM++*, J. Quant. Spectrosc. Ra., 167 (2015), pp. 40–52.

[54] S. P. GROTH, D. P. HEWETT, AND S. LANGDON, *Hybrid numerical–asymptotic approximation for high-frequency scattering by penetrable convex polygons*, IMA J. Appl. Math., 80 (2015), pp. 324–353.

[55] L. G. GUIMARÃES AND E. E. S. SAMPAIO, *A note on Snell's laws for electromagnetic plane waves in lossy media*, J. Quant. Spectrosc. Ra., 109 (2008), pp. 2124–2140.

[56] W. HACKBUSCH AND B. N. KHOROMSKIJ, *A sparse h-matrix arithmetic: general complexity estimates*, Journal of Computational and Applied Mathematics, 125 (2000), pp. 479–501.

[57] H. HADDAR, P. JOLY, AND H. M. NGUYEN, *Generalized impedance boundary conditions for scattering by strongly absorbing obstacles: The scalar case*, Math. Mod. Meth. Appl. Sci., 15 (2005), pp. 1273–1300.

[58] J. HARGREAVES, Y. W. LAM, S. LANGDON, AND D. P. HEWETT, *A high-frequency BEM for 3D acoustic scattering*, in The 22nd International Conference on Sound and Vibration, 2015.

[59] S. HAVEMANN AND A. J. BARAN, *Extension of t-matrix to scattering of electromagnetic plane waves by non-axisymmetric dielectric particles: application to hexagonal ice cylinders*, J. Quant. Spectrosc. Ra., 70 (2001), pp. 139–158.

[60] D. P. HEWETT, *Sound propagation in an urban environment*, PhD thesis, University of Oxford, 2010.

[61] ——, *Shadow boundary effects in hybrid numerical-asymptotic methods for high frequency scattering*, Euro. J. Appl. Math, 25 (2015), pp. 773–793.

[62] D. P. HEWETT, S. LANGDON, AND S. N. CHANDLER-WILDE, *A frequency-independent boundary element method for scattering by two-dimensional screens and apertures*, IMA J. Numer. Anal., 35 (2015), pp. 1698–1728.

[63] D. P. HEWETT, S. LANGDON, AND J. M. MELENK, *A high frequency hp boundary element method for scattering by convex polygons*, SIAM J. Numer. Anal., 51 (2013), pp. 629–653.

[64] A. J. HEYMSFIELD AND L. M. MILOSHEVICH, *Relative humidity and temperature influences on cirrus formation and evolution: Observations from wave clouds and FIRE II*, J. Atmos. Sci., 52 (1995), pp. 4302–4326.

[65] P. V. HOBBS, *Ice Physics*, Oxford: Clarendon Press, 1974.

[66] G. C. HSIAO AND L. XU, *A system of boundary integral equations for the transmission problem in acoustics*, Appl. Numer. Math., 61 (2011), pp. 1017–1029.

[67] D. HUYBRECHS, *Multiscale and hybrid methods for the solution of oscillatory integral equations*, PhD thesis, Department of Computer Science, Katholieke Universiteit Leuven, 2006.

[68] D. HUYBRECHS AND R. COOLS, *On generalized gaussian quadrature rules for singular and nearly singular integrals*, SIAM J. Numer. Anal., 47 (2009), pp. 719–739.

[69] D. HUYBRECHS AND S. OLVER, *Highly oscillatory quadrature*, in Highly Oscillatory Problems, Cambridge University Press, Cambridge, 2009, pp. 25–50.

[70] D. HUYBRECHS AND S. VANDEWALLE, *The efficient evaluation of highly oscillatory integrals in BEM by analytic continuation*, in Proceedings of the 5th UK Conference on Boundary Integral Methods (Chen, K., ed.), Liverpool University Press, Liverpool, 2005, pp. 20–31.

[71] ———, *A sparse discretization for integral equation formulations of high frequency scattering problems*, SIAM J. Sci. Comput., 29 (2007), pp. 2305–2328.

[72] A. ISERLES, S. P. NØRSETT, AND S. OLVER, *Highly oscillatory quadrature: The story so far*, in Numerical Mathematics and Advanced Applications, Springer, 2006, pp. 97–118.

[73] J. D. JACKSON, *Classical Electrodynamics*, vol. 3, John Wiley & Sons, Inc, 1962.

[74] G. L. JAMES, *Geometrical Theory of Diffraction for Electromagnetic Waves*, vol. 1, Peter Peregrinus Ltd, 1986.

[75] J.-M. JIN, *The Finite Element Method in Electromagnetics*, John Wiley & Sons, 2014.

[76] D. S. JONES, *The Theory of Electromagnetism*, Macmillan, 1964.

[77] C.-S. JOO, J.-W. RA, AND S.-Y. SHIN, *Scattering by right angle dielectric wedge*, IEEE T. Antenn. Propag., 32 (1984), pp. 61–69.

[78] F. M. KAHNERT, *Numerical methods in electromagnetic scattering theory*, J. Quant. Spectrosc. Ra., 79 (2003), pp. 775–824.

[79] J. B. KELLER, *Geometrical theory of diffraction*, J. Opt. Soc. Amer., 52 (1962), pp. 116–130.

[80] J. KEPLER, *The Six-cornered Snowflake*, Paul Dry Books, 2010.

[81] T. KIM, *Asymptotic and numerical methods for high-frequency scattering problems*, PhD thesis, University of Bath, 2012.

[82] R. KITTAPPA AND R. E. KLEINMAN, *Acoustic scattering by penetrable homogeneous objects*, J. Math. Phys., 16 (1975), pp. 421–432.

[83] R. E. KLEINMAN AND P. A. MARTIN, *On single integral equations for the transmission problem of acoustics*, SIAM J. Appl. Math., 48 (1988), pp. 307–325.

[84] ———, *On single integral equations for the transmission problem of acoustics*, SIAM J. Appl. Math., 48 (1988), pp. 307–325.

[85] A. A. KOCHANOVSKIJ, *Light Scattering Reviews: Single and Multiple Light Scattering*, Springer, 2006.

[86] R. G. KOUYOUMJIAN AND P. H. PATHAK, *A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface*, P. IEEE, 62 (1974), pp. 1448–1461.

[87] S. V. KULKARNI AND S. A. KHAPARDE, *Transformer Engineering: Design and Practice*, vol. 25, CRC Press, 2004.

[88] A. R. LALIENA, M.-L. RAPÚN, AND F.-J. SAYAS, *Symmetric boundary integral formulations for helmholtz transmission problems*, Appl. Numer. Math., 59 (2009), pp. 2814–2823.

[89] R. P. LAWSON, B. PILSON, B. BAKER, Q. MO, E. JENSEN, L. PFISTER, AND P. BUI, *Aircraft measurements of microphysical properties of subvisible cirrus in the tropical tropopause layer*, Atmos. Chem. Phys., 8 (2008), pp. 1609–1620.

[90] L. LEWIN AND I. SREENIVASIAH, *Diffraction by a dielectric wedge.* Scientific report no. 47, 1979. Department of Electrical Engineering, University of Colorado.

[91] K. G. LIBBRECHT, *SnowCrystals.com*, Aug. 2015.

[92] K. N. LIOU, *Influence of cirrus clouds on weather and climate processes: A global perspective*, Mon. Weather Rev., 114 (1986), pp. 1167–1199.

[93] H. M. MACDONALD, *Electric Waves*, Cambridge University Press, 1902.

[94] A. MACKE, *Scattering of light by polyhedral ice crystals*, Appl. Optics, 32 (1993), pp. 2780–2788.

[95] A. MACKE, J. MUELLER, AND E. RASCHKE, *Single scattering properties of atmospheric ice crystals*, J. Atmos. Sci., 53 (1996), pp. 2813–2825.

[96] A. I. MAHAN AND C. V. BITTERLI, *Reflection and transmission of plane unbounded electromagnetic waves at an absorbing–nonabsorbing interface with numerical calculations for an ocean–air interface*, Appl. optics, 20 (1981), pp. 3345–3359.

[97] G. D. MALIUZHINETS, *Excitation, reflection and emission of surface waves from a wedge with given face impedances*, Sov. Phys. Dokl., 3 (1958), pp. 572–755.

[98] S. MARBURG, *Discretization requirements: how many elements per wavelength are necessary?*, in Computational Acoustics of Noise Propagation in Fluids-Finite and Boundary Element Methods, Springer, 2008, pp. 309–332.

[99] E. MARMOLEJO-OLEA, I. MITREA, M. MITREA, AND Q. SHI, *Transmission boundary problems for Dirac operators on Lipschitz domains and applications to Maxwell's and Helmholtz's equations*, Trans. Amer. Math. Soc., 364 (2012), pp. 4369–4424.

[100] P. A. MARTIN, *On the null-field equations for the exterior problems of acoustics*, Q. J. Mech. Appl. Math., 33 (1980), pp. 385–396.

[101] P. A. MARTIN AND P. OLA, *Boundary integral equations for the scattering of electromagnetic waves by a homogeneous dielectric obstacle*, Proc. Roy. Soc. Edinburgh, 123A (1993), pp. 185–208.

[102] J. R. MAUTZ AND R. F. HARRINGTON, *Electromagnetic scattering from a homogeneous material body of revolution*, tech. rep., Department of Electrical and Computer Engineering, Syracuse University, 11 1977.

[103] W. MCLEAN, *Strongly Elliptic Systems and Boundary Integral Equations*, Cambridge University Press, 2000.

[104] E. MEISTER, F. PENZEL, F. O. SPECK, AND F. S. TEIXEIRA, *Two canonical wedge problems for the Helmholtz equation*, Math. Methods Appl. Sci., 17 (1994), pp. 877–899.

[105] M. I. MISHCHENKO, L. D. TRAVIS, AND A. A. LACIS, *Scattering, Absorption, and Emission of Light by Small Particles*, Cambridge University Press, 2002.

[106] N. MORITA, *Resonant solutions involved in the integral equation approach to scattering from conducting and dielectric cylinders*, IEEE T. Antenn. Propag., 27 (1979), pp. 869–871.

[107] K. MUINONEN, *Scattering of light by crystals: a modified kirchhoff approximation*, Appl. Optics, 28 (1989), pp. 3044–3050.

[108] ——, *Scattering of light by crystals: a modified kirchhoff approximation*, Appl. Optics, 28 (1989), pp. 3044–3050.

[109] K. MUINONEN, K. LUMME, J. PELTONIEMI, AND W. M. IRVINE, *Light scattering by randomly oriented crystals*, Appl. Optics, 28 (1989), pp. 3051–3060.

[110] B. J. MURRAY, C. G. SALZMANN, A. J. HEYMSFIELD, S. DOBBIE, R. R. NEELY-III, AND C. J. COX, *Trigonal ice crystals in earth's atmosphere*, Bulletin of the American Meteorological Society, (2014).

[111] J. C. NÉDÉLEC, *Acoustic and Electromagnetic Equations: Integral Representations for Harmonic Problems*, Springer, New York, 2001.

[112] X.-C. NIE, L.-W. LI, N. YUAN, T. S. YEO, AND Y.-B. GAN, *Precorrected-fft solution of the volume integral equation for 3-d inhomogeneous dielectric objects*, IEEE Transactions on Antennas and Propagation, 53 (2005), pp. 313–320.

[113] J. R. OCKENDON, S. HOWISON, A. LACEY, AND A. MOVCHAN, *Applied Partial Differential Equations*, Oxford University Press, 2003.

[114] F. W. J. OLVER, D. W. LOZIER, R. F. BOISVERT, AND C. W. CLARK, eds., *NIST Handbook of Mathematical Functions*, Cambridge University Press, New York, NY, 2010.

[115] E. PAROLIN, *A hybrid numerical-asymptotic boundary element method for high-frequency wave scattering*, Master's thesis, University of Oxford, 2015.

[116] L. PINCHERLE, *Refraction of plane non-uniform electromagnetic waves between absorbing media*, Physical Review, 72 (1947), p. 232.

[117] J. RADLOW, *Diffraction by a right-angled dielectric wedge*, Int. J. Eng. Sci., 2 (1964), pp. 275–290.

[118] M.-L. RAPÚN AND F.-J. SAYAS, *Mixed boundary integral methods or helmholtz transmission problems*, J. Comput. Appl. Math., 214 (2008), pp. 238–258.

[119] P.-A. RAVIART AND J.-M. THOMAS, *A mixed finite element method for 2-nd order elliptic problems*, in Mathematical aspects of finite element methods, Springer, 1977, pp. 292–315.

[120] A. D. RAWLINS, *Diffraction by a dielectric wedge*, IMA J. Appl. Math., 19 (1977), pp. 261–279.

[121] ——, *Diffraction by, or diffusion into, a penetrable wedge*, Proc. R. Soc. Lond. A, 455 (1999), pp. 2655–2686.

[122] R. ROSHDI, *A pioneer in anaclastics: Ibn Sahl on burning mirrors and lenses*, JSTOR, 81 (1990), pp. 464–491.

[123] K. SASSEN, Z. WANG, AND D. LIU, *Global distribution of cirrus clouds from CloudSat/Cloud-Aerosol lidar and infrared pathfinder satellite observations (CALIPSO) measurements*, J. Geophys. Res.-Atmos., 113 (2008).

[124] W. ŚMIGAJ, T. BETCKE, S. ARRIDGE, J. PHILLIPS, AND M. SCHWEIGER, *Solving boundary integral problems with BEM++*, ACM T. Math. Software, (2015).

[125] H. R. SMITH, A. WEBB, P. CONNOLLY, A. J. BARAN, A. R. D. SMEDLEY, AND E. HESSE, *Cloud chamber laboratory investigations into the scattering properties of hollow ice particles*, J. Quant. Spectrosc. Ra., 157 (2015), pp. 106–118.

[126] A. SOMMERFELD, *Mathematische Theorie der Diffraction*, Math. Ann., 47 (1896), pp. 317–374.

[127] E. A. SPENCE, S. N. CHANDLER-WILDE, I. G. GRAHAM, AND V. P. SMYSHLYAEV, *A new frequency-uniform coercive boundary integral equation for acoustic scattering*, Communications on Pure and Applied Mathematics, 64 (2011), pp. 1384–1415.

[128] E. A. SPENCE AND A. FOKAS, *A new transform method i: domain-dependent fundamental solutions and integral representations*, in Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 466, The Royal Society, 2010, pp. 2259–2281.

[129] E. A. SPENCE, I. V. KAMOTSKI, AND V. P. SMYSHLYAEV, *Coercivity of Combined Boundary Integral Equations in High-Frequency Scattering*, Communications on Pure and Applied Mathematics, 68 (2015), pp. 1587–1639.

[130] J. A. STRATTON, *Electromagnetic Theory*, Wiley-Interscience, 2007.

[131] W. SUN, Q. FU, AND Z. CHEN, *Finite-difference time-domain solution of light scattering by dielectric particles with a perfectly matched layer absorbing boundary condition*, Appl. Optics, 38 (1999), pp. 3141–3151.

[132] A. TAFLOVE AND S. C. HAGNESS, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, Artech House, INC., (2005).

[133] R. H. TORRES AND G. V. WELLAND, *The Helmholtz equation and transmission problems with Lipschitz interfaces*, Indiana Univ. Math. J., 42 (1993), pp. 1457–1486.

[134] L. N. TREFETHEN AND D. BAU, *Numerical Linear Algebra*, SIAM, 1997.

[135] A. TWIGGER, *Boundary Element Methods for High Frequency Scattering*, PhD thesis, University of Reading, 2012.

[136] P. Y. UFIMTSEV, *Fundamentals of the Physical Theory of Diffraction*, John Wiley & Sons, 2007.

[137] R. J. UNCLES, *Numerical solution of acoustic scattering problems at intermediate frequencies*, J. Acoust. Soc. Am., 60 (1976), p. 266.

[138] A. M. VOGELMANN AND T. P. ACKERMAN, *Relating cirrus cloud properties to observed fluxes: A critical assessment*, Journal of the atmospheric sciences, 52 (1995), pp. 4285–4301.

[139] S. G. WARREN, *Optical constants of ice from the ultraviolet to the microwave*, Appl. Optics, 23 (1984), pp. 1206–1225.

[140] S. G. WARREN AND R. E. BRANDT, *Optical constants of ice from the ultraviolet to the microwave: A revised compilation*, J. Geophys. Res.-Atmos., 113 (2008).

[141] P. C. WATERMAN, *Matrix formulation of electromagnetic scattering*, Proc. IEEE, 53 (1965), pp. 805–812.

[142] M. WENDISCH AND P. YANG, *Theory of atmospheric radiative transfer*, John Wiley & Sons, 2012.

[143] P. WENDLING, R. WENDLING, AND H. K. WEICKMANN, *Scattering of solar radiation by hexagonal ice crystals*, Appl. Optics, 18 (1979), pp. 2663–2671.

[144] T. WRIEDT, *A review of elastic light scattering theories*, Part. Part. Syst. Char., 15 (1998), pp. 67–74.

[145] P. YANG AND K. N. LIOU, *Light scattering by hexagonal ice crystals: comparison of finite-difference time domain and geometric optics models*, J. Opt. Soc. Amer. A, 12 (1995), pp. 162–176.

[146] ——, *Finite-difference time domain method for light scattering by small ice crystals in three-dimensional space*, J. Opt. Soc. Am. A, 13 (1996), pp. 2072–2085.

[147] ——, *Light scattering by hexagonal ice crystals: solutions by a ray-by-ray integration algorithm*, J. Opt. Soc. Am. A., 14 (1997), pp. 2278–2289.

[148] P. YANG AND K. N. LIOU, *Effective refractive index for determining ray propagation in an absorbing dielectric particle*, J. Quant. Spectrosc. Ra., 110 (2009), pp. 300–306.

[149] P. YANG AND K. N. LIOU, *An exact geometric-optics approach for computing the optical properties of large absorbing particles*, J. Quant. Spectrosc. Ra., 110 (2009), pp. 1162–1177.

[150] K. YEE, *Numerical solution of initial boundary value problems involving Maxwells equations in isotropic media*, IEEE Trans. Antenn. Propag., 14 (1966), pp. 302–307.

[151] M. A. YURKIN AND A. G. HOEKSTRA, *The discrete dipole approximation: an overview and recent developments*, J. Quant. Spectrosc. Ra., 106 (2007), pp. 558–589.