Approximation by dominant wave directions in plane wave methods

T. Betcke, J. Phillips

University College London, Gower Street, WC1E 6BT, UK

Abstract

Plane wave methods have become an established tool for the solution of Helmholtz problems in homogeneous media. The idea is to approximate the solution in each element with a linear combination of plane waves, which are roughly equally spaced in all directions. The main advantage of plane wave methods is that they require significantly fewer degrees of freedom per unknown than standard finite elements. However, for many wave problems there are only a few dominant wave directions, which can be found using ray tracing or other high-frequency methods. Based on arguments from high-frequency asymptotics we show that dominant plane waves can be a suitable approximation basis if multiplied (modulated) by small degree polynomials. We explore this approach for two example problems and compare its performance against standard equispaced plane wave basis sets. Finally, we present an example with smoothly varying speed of sound, which demonstrates that for such problems approximations based on polynomially modulated dominant wave directions can far outperform standard plane wave methods, which are not well suited for handling problems in varying media.

Keywords: plane wave methods, ray tracing, Helmholtz

1. Introduction

In recent years plane wave methods (see e.g. [1, 2, 3, 4]) have become a widely used tool for the solution of Helmholtz problems of the form

$$\Delta u + \left(\frac{\omega}{c}\right)^2 u = 0, \tag{1}$$

where ω is the frequency and c is the speed of sound (we will also use the spatial wavenumber $k := \omega/c$). Typically, plane wave methods are applied only if c is constant or at least piecewise constant. Later in this paper we will also consider the case that $c = c(x)$ is varying across the domain. The idea of plane wave methods for constant or piecewise constant c is to approximate the solution u in an element K by a linear combination of plane waves of the form

$$u(x) \approx \sum_{j=0}^{N-1} \alpha_j e^{i\frac{\omega}{c} d_j \cdot x}, \tag{2}$$

where $\alpha_j \in \mathbb{C}$ and the $d_j \in \mathbb{R}^d$, $\|d_j\| = 1$, are direction vectors in $d = 2$ or $d = 3$ dimensions. In 2-d a typical choice for the direction vectors d_j is

$$d_j = \begin{bmatrix} \cos\theta_j, & \sin\theta_j \end{bmatrix}^T, \quad \theta_j = \frac{2\pi j}{N}. \tag{3}$$

In 3-d a suitable choice of directions is for example described in [5]. Plane wave methods often require substantially fewer degrees of freedom for a given accuracy than standard finite element methods. The reason is that the basis

functions are already solutions of the underlying Helmholtz equation and therefore have the right oscillatory behavior. This is one example of a wider class of methods, called Trefftz methods, which approximate solutions of PDE problems using a basis composed of particular solutions of the homogeneous partial differential operator. Other frequently used choices of basis functions include Fourier-Bessel functions [6, 7, 8] and fundamental solutions [9, 10].

One potential advantage of plane wave methods is the freedom in the selection of the plane wave directions. Standard approximation estimates for plane waves, such as those in [5], necessarily assume that the directions are approximately equally spaced on the unit sphere. However, in many wave problems, such as scattering from convex obstacles, there are only few dominant wave directions, and it seems natural to try to exploit this fact. One way to determine dominant directions is to use ray tracing, which is particularly simple for problems in homogeneous media.

Assume that a Helmholtz solution has a single dominant wave direction $\hat{d}$ in an element K. Then a naive approach would be to just approximate the solution u in K by $\tilde{u} := \alpha e^{i\frac{\omega}{c}\hat{d}\cdot x}$ for some constant α. But this only gives an acceptable solution if u itself is very close to a plane wave. A more general approach, based on high-frequency arguments, is to use functions of the form

$$\sum_{j=0}^{M-1} p_j(x) e^{i\frac{\omega}{c}\hat{d}_j \cdot x}, \tag{4}$$

polynomially modulated plane waves, where the p_j are small degree polynomials, and the $\hat{d}_j$ are the dominant wave directions obtained from ray tracing. Related types of basis functions have already been proposed in the context of high-frequency finite element methods [11]. The idea in high-frequency finite elements is to exploit that in the limit $\omega \to \infty$ a Helmholtz solution u has the asymptotic representation

$$u = \sum_j A_j(x) e^{i\omega\phi_j(x)} + O\left(\frac{1}{\omega}\right), \tag{5}$$

where the A_j are amplitude functions and the ϕ_j phase functions. The amplitude functions A_j are slowly oscillating and can therefore be efficiently computed once the ϕ_j are known. However, determining the ϕ_j is often difficult, especially in problems with varying speed of sound c or in the presence of complex geometries. Here, we only assume that approximations to the dominant wave directions $\hat{d}_j$ are available.

For the assembly we use a modified formulation of the Plane Wave Discontinuous Galerkin Method (PWDG) introduced in [12], which keeps the volume terms. This discontinuous Galerkin ansatz makes it easy to include different dominant directions in different elements.

The paper is organised as follows. In Section 2 we give a brief overview of some concepts from high-frequency asymptotics, which motivate the concepts in this paper. Then in Section 3 we demonstrate the approximation properties of polynomially modulated dominant plane waves in comparision to standard equispaced plane waves. A discontinuous Galerkin approach to assemble the matrices for a computation on a mesh is presented in Section 4. This is used in Section 5 and 6 to investigate two example problems. The scattering problem in Section 6 also contains diffracted waves and we propose one approach of incorporating them. In Section 7 we present an example involving a problem with varying speed of sound. We conclude the paper with remarks in Section 8.

2. Elements of high-frequency asymptotics

In this section we briefly summarize some results from high-frequency asymptotics, which will be needed in this paper. A more detailed introduction is given in [13].

We assume that the solution u of (1) can be represented as

$$u(x) = e^{i\omega\phi(x)} \sum_{\ell=0}^{\infty} a_\ell(x)(i\omega)^{-\ell}.$$

Here, $\phi(x)$ is an unknown phase function and the a_ℓ are unknown amplitude functions. Plugging into (1) and matching leading order coefficients in ω gives

$$|\nabla\phi(x)| = \frac{1}{c(x)}, \tag{6}$$
$$\Delta\phi A + 2\nabla\phi\nabla A = 0. \tag{7}$$

2

Equation (6) is an eikonal equation for the phase function ϕ and (7) is a transport equation for the leading order amplitude function $A(x) := a_0(x)$. After solving for ϕ and A, an approximation to the Helmholtz solution u, is

$$u(x) = A(x)e^{i\omega\phi(x)} + O\left(\frac{1}{\omega}\right).$$

For multiple scattering configurations the solution for the eikonal equation is not unique, giving rise to the representation in (5).

Now define the Hamiltonion $H(x, p) := |p|c(x)$. Obviously, from (6) it follows that $H(x, \nabla\phi(x)) \equiv 1$. The Hamiltonian is constant along the bicharacteristic curves, $(x(t), p(t))$ defined by

$$\frac{d}{dt}x(t) = \nabla_p H(x, p) = c(x)\frac{p}{|p|}, \tag{8}$$

$$\frac{d}{dt}p(t) = -\nabla_x H(x, p) = -|p|\nabla c(x). \tag{9}$$

By rescaling p, we have $H(x(t), p(t)) = 1$ and a short calculation, given in [13], shows that the pair $(x(t), \nabla\phi(x(t)))$ satisfies (8)-(9). So by uniqueness we have $p(t) = \nabla\phi(x(t))$.

Defining $\eta(x) = c(x)^{-1}$ and rescaling (8)-(9) such that $|p|c(x) = 1$ gives the *ray equations*

$$\frac{d}{dt}x(t) = \frac{1}{\eta^2}p, \tag{10}$$

$$\frac{d}{dt}p(t) = \frac{\nabla\eta}{\eta} \tag{11}$$

with initial conditions $x(0) = x_0$, $p(0) = p_0$ under the constraint that $|p_0| = \eta(x_0)$. Ray tracing consists of solving this system of ODEs. The solutions, $x(t)$, are called rays and p is called the slowness vector. For homogeneous media η is constant. It follows that $p(t)$ is constant along the rays $x(t)$, which are just straight lines. Also, since $p = \nabla\phi(x)$, from (10) it follows that the rays are orthogonal to the level sets of ϕ.

We now define what we mean by a dominant wave direction in an element K.

Definition 1 (Dominant wave direction). *Let ϕ be a solution of the eikonal equation* (6)*, which is differentiable in K. The dominant wave direction in K associated with ϕ is defined as the average*

$$d := \int_K \nabla\phi(x)dx.$$

If the solution of the eikonal equation in K is non unique, as it is the case in multiple scattering problems, each branch of the multi-valued solution to the eikonal solution will lead to a different dominant direction. This is natural for the case that the Helmholtz solution u in K is composed of waves travelling in different directions.

In practice we assume that a wave is only slowly changing in an element. Hence, a good approximation to a dominant direction is to set $d := \nabla\phi(x_0)$ for some $x_0 \in K$. If $x_0 = x(t)$ for some t, then $d = p(t)$. Hence, we need to do ray tracing in such a way that the element K is hit by a ray. For constant media problems this is usually not difficult. For problems in inhomogeneous media rays may diverge exponentially in which case, for example, wavefront tracking methods can be used [14].

3. Approximating a Helmholtz solution by a dominant plane wave

Consider a sound source at 0 of the form $u(r) = H_0^{(1)}(k|x|)$. If we are sufficiently far from the origin, this function locally appears to be a plane wave. We now consider the best approximation of u in the square $[1, 2 + h] \times [2, 2 + h]$, where $h = 0.2$ by a plane wave of the form $\alpha e^{ikd\cdot x}$, where $d = \begin{bmatrix} \cos\theta, & \sin\theta \end{bmatrix}^T$. As wavenumber we choose $k = 30$. Hence, for the wavelength λ we have $\lambda = \frac{2\pi}{k} \approx h$. The situation is depicted in the left plot of Figure 1. The right plot of Figure 1 shows the relative L^2 approximation error in dependence on the angle θ. The 1% error level is shown by a dashed line. The best possible error is about 3% with only one degree of freedom. Naturally, if we move further
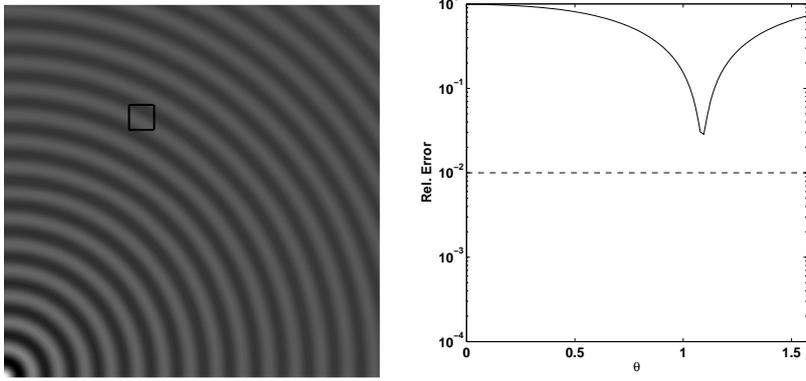
Figure 1: Left: A sound source with wavenumber $k = 30$. In the square we approximate the field by a single plane wave. The right plot shows the relative approximation error in the square in dependence of the angle θ of the plane wave.

away from the source the wave appears more like a plane wave and we expect the error to be smaller. This artificial example demonstrates that we can approximate a non-trivial Helmholtz solution locally by using a dominant plane wave direction. However, we need to correct the error between the dominant plane wave and the exact solution. In the high-frequency limit we assumed a representation of the form $u(x) \sim A(x)e^{i\omega\phi(x)}$ as $\omega \to \infty$ with a smooth amplitude function that is the solution of the transport equation (7). For finite frequencies this is still a valid assumption. If we make the ansatz $u(x) = \tilde{A}(x)e^{i\omega\phi(x)}$ we have, by substituting into the Helmholtz equation (1) the relationship [15],

$$
\begin{aligned}
0 &= \frac{1}{\omega}\Delta\tilde{A} + 2i\nabla\tilde{A}\cdot\nabla\phi + i\tilde{A}\Delta\phi - \omega\left(|\nabla\phi|^2 - \eta^2\right)\tilde{A} \\
&= \frac{1}{\omega}\Delta\tilde{A} + 2i\nabla\tilde{A}\cdot\nabla\phi + i\tilde{A}\Delta\phi.
\end{aligned}
\tag{12}
$$

As $\omega \to \infty$ we recover the transport equation (7). Hence, the diffusion term $\frac{1}{\omega}\Delta\tilde{A}$ is a finite frequency correction term. If the high-frequency approximation $A(x)e^{i\omega\phi(x)}$ is a good approximate model for the wave then we can expect that also the finite frequency amplitude function $\tilde{A}$ is only mildly oscillatory since all oscillations are contained in $e^{i\omega\phi(x)}$ if ϕ is the correct phase function.

In the model problem we approximated the phase contribution $e^{i\phi(x)}$ by a plane wave. But the amplitude contribution $\tilde{A}$ was only approximated by a constant α in the small square. To capture the amplitude contribution better we approximate u by a function of the form $\tilde{u} := p(x)e^{i\frac{\omega}{c}d\cdot x}$. Let $\|\cdot\|_{0,K}$ be the $L^2(K)$ norm and suppose that diam $K \le h$. Furthermore, without loss of generality we can assume that 0 is in the interior of K and that $\phi(0) = 0$. This is obtained by rewriting $A(x)e^{i\omega\phi(x)} = A(x)e^{i\omega\phi(x_0)}e^{i\omega(\phi(x_0+x)-\phi(x_0))}$ for some $x_0 \in K$. Then

$$
\begin{aligned}
\|u - \tilde{u}\|_{0,K} &= \|p(x)e^{i\frac{\omega}{c}d\cdot x} - \tilde{A}(x)e^{i\omega\phi(x)}\|_{0,K} \\
&\le \|e^{i\frac{\omega}{c}d\cdot x}\|_{0,K}\left(\|p(x) - \tilde{A}(x)\|_{0,K} + \|\tilde{A}\|_{0,K}\|e^{i\omega(\phi(x)-\frac{d}{c}\cdot x)} - 1\|_{0,K}\right)
\end{aligned}
\tag{13}
$$

Now, $e^{i\omega(\phi(x)-\frac{d}{c}\cdot x)} - 1 \approx i\omega(\phi(x) - \frac{d}{c}\cdot x)$ if $\omega\|\phi(x) - \frac{d}{c}\cdot x\|_{0,K}$ is sufficiently small. Hence, from (13) it follows that

$$
\|u - \tilde{u}\|_{0,K} \lesssim \|e^{i\frac{\omega}{c}d\cdot x}\|_{0,K}\left(\|p(x) - \tilde{A}(x)\|_{0,K} + \omega\|\tilde{A}\|_{0,K}\|\phi(x) - \frac{d}{c}\cdot x\|_{0,K}\right).
\tag{14}
$$

There are two error sources, namely the error of approximating $\tilde{A}$ by a polynomial p in the element K and the error of approximating ϕ by a linear function $d \cdot x$ in K. While the first error can be controled by the polynomial degree of the finite element space the second error depends on how well ϕ can be approximated by a linear function in the element K. Assume that ϕ is sufficiently smooth. Then

$$
\phi(x) - \frac{d}{c}\cdot x = \left[\nabla\phi(0) - \frac{d}{c}\right]\cdot x + O(h^2)
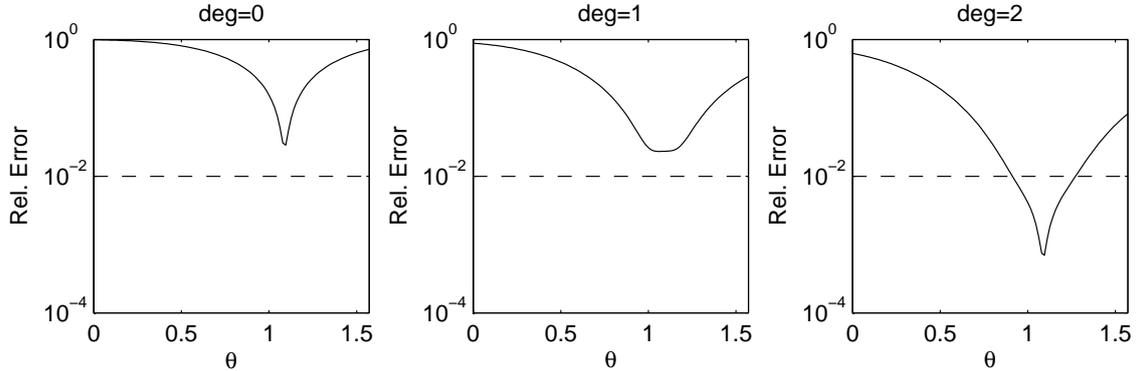$$

Figure 2: Approximating the wave in the small square shown in Figure 1. This time the approximating wave is modulated with polynomials of degree 0 (left plot, same as the approximation in Figure 1), 1 (middle plot), and 2 (right plot).

since $\|x\| \leq h$ in K. Hence, by denoting the angle between $\nabla\phi(0)$ and d as θ and assuming that $\theta \in [0, \pi/2]$, up to quadratic terms in h we have

$$\|\phi(x) - d \cdot x\| \lessapprox h\|\nabla\phi(0) - \frac{d}{c}\| \leq \sqrt{2}\frac{h}{c}\sin\theta.$$

Combining this estimate with (13) results in

$$\|u - \tilde{u}\|_{0,K} \lessapprox \|e^{i\omega d \cdot x}\|_{0,K}\left(\|p(x) - \tilde{A}(x)\|_{0,K} + \sqrt{2}\|\tilde{A}\|_{0,K}\frac{h\omega}{c}\sin\theta\right). \tag{15}$$

Assuming that $\tilde{A}$ is sufficently smooth and not oscillatory then a low degree polynomial p is sufficient to approximate $\tilde{A}$ on large elements. The main error contribution is therefore given by the second term, requiring that $\frac{h\omega}{c}\sin\theta \ll 1$. In standard finite elements the requirement is that $\frac{h\omega}{c} \ll 1$. Hence, if we have a good approximation to the dominant wave directions we obtain a significant improvement to standard finite element methods. Indeed, if $\nabla\phi(0) = \frac{d}{c}$ then up to quadratic terms in h the oscillations in K are completely canceled out, allowing for significantly larger element sizes than with standard approaches.

This analysis ignores that a sufficiently high degree polynomial p will also approximate the oscillatory phase factor in an element well. However, the goal here is a different one. We want to have large elements with small degree polynomials that are just sufficient to approximate the amplitude contribution well.

Figure 2 shows a similar experiment as the right plot of Figure 1, but this time we approximate with functions of the form $p(x)e^{i\frac{\omega}{c}d \cdot x}$. Certainly, the situation is artificial, but it demonstrates that optimally chosen directions can be efficient in approximating Helmholtz solutions if the amplitude contribution is taken into account by modulating with polynomial terms. We also note that it is not required to choose the best direction in an element. In the right plot of Figure 2 we see that with polynomials of degree 2, there is a range of θ for which the error level remains below 1%. Hence, it ought to be sufficient to have approximate dominant directions in each element.

In Figure 3 we show the approximation error with a basis of equidistant plane wave directions as in (2). In this example we need at least 11 directions to achieve a relative error of less than 1%. The same level of accuracy is achieved with only 6 basis functions in the modulated basis used for the right plot of Figure 2. This indicates that modulated plane waves can be an efficient alternative to equidistant plane waves if there is only one dominant wave direction in each element.

4. A discontinuous Galerkin formulation

Various approaches have been proposed for the assembly of a discrete system using plane wave type bases, such as the PUFEM [16, 17], the Ultra Weak Variational Formulation [1, 2] or Lagrange multiplier methods [18]. The UWVF
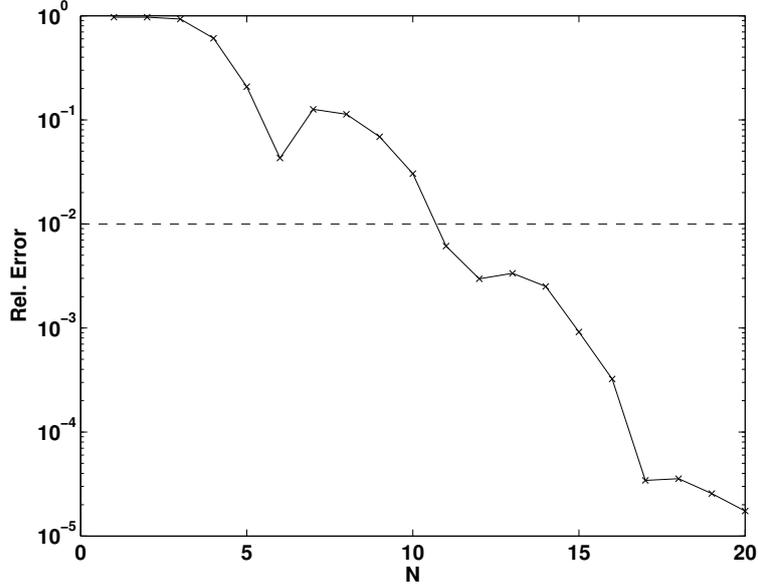
Figure 3: Error of approximating the wave in the small square from the left plot of Figure 1 by a linear combination of plane waves of the form (2).

and Lagrange multiplier methods make use of the fact that the basis functions satisfy the exact Helmholtz equation (1) in the interior of each element. Hence, the formulation only needs to match up the inter-element continuity conditions and the boundary conditions. Modulated plane waves in general do not satisfy the Helmholtz equation. Hence, we also need to assemble volume terms. But this can easily be accommodated. In this section we briefly demonstrate this at the example of the Plane Wave Discontinuous Galerkin Method (PWDG), a generalisation of the UWVF, introduced by Gittelson, Hiptmair and Perugia [12]. This method will be used for the numerical experiments presented later in this paper.

For the derivation in this section we rewrite (1) in the form

$$-\Delta u - k^2 \mu^2(x) u(x) = 0, \tag{16}$$

where $k = \omega/c_0$ for some reference speed of sound c_0 and $\mu(x) = c_0/c(x)$.

Consider a triangularization $\mathcal{T}_h$ of the computational domain Ω and let $K \in \mathcal{T}_h$ be an element of the triangularisation. Multiplying (1) by a test function v in element K and integrating by parts gives

$$\int_K \nabla u \cdot \overline{\nabla v} dV - k^2 \int_K \mu^2 u \overline{v} dV - \int_{\partial K} \nabla u \cdot n \overline{v} dS = 0.$$

A further integration by parts yields

$$\int_K \overline{(-\Delta v - k^2 \mu^2) v} u dV + \int_{\partial K} u \cdot \overline{\nabla v \cdot n} dS - \int_{\partial K} \nabla u \cdot n \overline{v} dS = 0. \tag{17}$$

Define $\sigma := \frac{1}{ik} \nabla u$. If μ is constant in K and the test function v is a Helmholtz solution with respect to the wavenumber $k\mu$ in K, that is $-\Delta v - k^2 \mu^2 v = 0$, then

$$\int_{\partial K} u \cdot \overline{\nabla v \cdot n} dS - ik \int_{\partial K} \sigma \cdot n \overline{v} dS = 0. \tag{18}$$

The standard discontinuous Galerkin approach is now to approximate u and σ on ∂K by fluxes $\hat{u}_h$ and $\hat{\sigma}_h$, which are linear combinations of jumps and averages of function values and normal derivatives in the respective neighboring

6

elements. In particular, in [12] the choice

$$\hat{\sigma}_h := \frac{1}{ik} \{\!\{\nabla_h u_h\}\!\} - \alpha \, [\![u_h]\!] - \frac{\gamma}{ik} \, [\![\nabla_h u_h]\!]$$

$$\hat{u}_h := \{\!\{u_h\}\!\} + \gamma \cdot [\![u_h]\!] - \frac{\beta}{ik} \, [\![\nabla_h u_h]\!] \tag{19}$$

was proposed. The index h denotes that the functions u_h and σ_h and the operator ∇_h are to be understood elementwise. For a scalar v the jump $[\![v]\!]$ and average $\{\!\{v\}\!\}$ on a boundary Γ between two elements K^- and K^+ is defined by

$$\{\!\{v\}\!\} := \frac{1}{2}\left(v^- + v^+\right), \quad [\![v]\!] := v^- n^- + v^+ n^+.$$

For a vector quantity σ we have

$$\{\!\{\sigma\}\!\} := \frac{1}{2}\left(\sigma^- + \sigma^+\right), \quad [\![\sigma]\!] := \sigma^- \cdot n^- + \sigma^+ \cdot n^+,$$

where n^- and n^+ are the outgoing normal directions with respect to K^- and K^+. The PWDG discretisation is now given by (18) with flux definitions in (19). Suitable choices for the parameters α, β and γ and flux definitions for impedance boundary conditions are discussed in [12]. In order to incorporate basis and test functions, which do not satisfy the Helmholtz equation (1) we simply reverse the second integration by parts in (17) and obtain, after including the flux terms, the formulation (see also [12, eq. 2.5])

$$\int_K \nabla u_h \cdot \overline{\nabla v_h} dV - k^2 \int_K \mu^2 u_h \overline{v_h} dV + \int_{\partial K} (\hat{u}_h - u_h) \cdot \overline{\nabla v_h \cdot n} dS - ik \int_{\partial K} \hat{\sigma}_h \cdot n \overline{v_h} dS = 0. \tag{20}$$

Hence, the additional volume term is given by a block-diagonal matrix whose entries are defined by the weak form

$$w(u_h, v_h) := \int_K \nabla u_h \cdot \overline{\nabla v_h} dV - k^2 \int_K \mu^2 u \overline{v_h} dV - \int_{\partial K} u_h \cdot \overline{\nabla v_h \cdot n} dS.$$

For the flux parameters α, β, and γ we choose the following values.

$$\alpha = \frac{ap^2}{kh}, \quad \beta = \frac{bkh}{p}, \quad \gamma = 0.$$

Here, p is the polynomial degree of the modulated plane wave basis. Without p, corresponding to no polynomial modulation, these flux parameters were suggested in [12]. If also $\beta = 0$ then the formulation corresponds to an interior penalty DG (IP-DG) method [19]. For the IP-DG parameter α the scaling by p^2 is a standard choice (see e.g. [20]). The division by p in the β parameter can be justified by scaling arguments. The values $a, b > 0$ are constants. For simplicity we use $a = b = 1$.

5. A simple numerical example

To illustrate the behaviour of the modulated plane wave basis within the discontinuous Galerkin framework, we will approximate the Helmholtz solution due to a source at $(-0.2, -0.2)$ in the square $[0, 1]^2$ (see Figure 4). We use the formulation from the previous section and impose an impedance boundary condition consistent with the desired solution.

On an element K with centre c_K, the dominant direction is taken to be the unit vector along $d_K := c_K - (-0.2, -0.2)$. On each element, therefore, we approximate with functions from a basis set consisting of modulated plane waves of the form $p(x)e^{ik\mathbf{d}_K \cdot x}$, where $p(x)$ is a polynomial of degree p, whose basis functions are L^2-orthogonal polynomials on K. Note that a polynomial of degree p in 2-d is represented by $(p + 1)(p + 2)/2$ basis functions. For comparison, we will use the basis of N equispaced plane waves shown in (2), where the unit vectors, d_j are as defined in (3).

The relative errors for these choices of bases are shown for $k = 20$ and $k = 60$ in the top (plane waves) and middle (modulated waves) plots of Figure 5. Ignoring conditioning issues with plane waves if m is large, equispaced plane
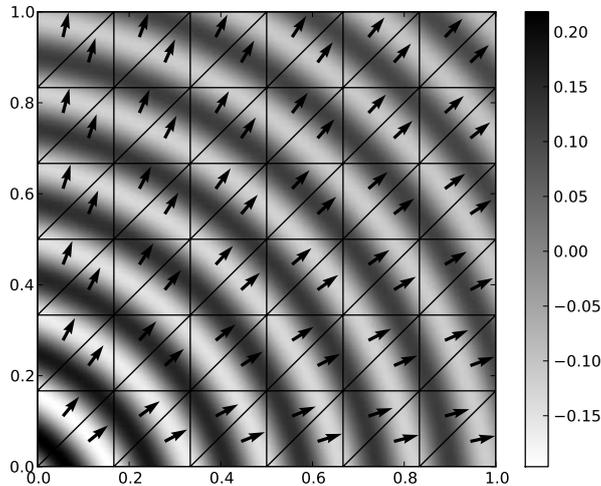
Figure 4: The real part of the Hankel function with origin at $(-0.2, -0.2)$, $g(x) := H_0(k |x + (0.2, 0.2)|)$ with ray-traced directions on each element.

waves have a better asymptotic error behavior compared to modulated bases with a similar number of unknowns. Here the effect becomes visible that polynomials of degree p need $O(p^2)$ basis functions to have a similar approximation order as $O(p)$ plane wave directions as $h \to 0$ [5].

But in the bottom plots of Figure 5 we see that the asymptotic behaviour is not the full story. The modulated plane wave basis is able to achieve a modest relative error (e.g. 10^{-2}), using significantly fewer degrees of freedom. The effect is even more pronounced when we let $k = 60$.

The deterioration of the error seen as $h \to 0$ for some of the plane wave basis sets is explained by the condition numbers in Table 1. Condition numbers for the modulated plane wave bases are given in Table 2. The condition numbers of the systems required to achieve a given error are broadly comparable. For example, at $h = 0.167$, the basis composed of plane waves modulated by polynomials of degree 2 achieves a relative error of 1.4% (it goes up to 2.1% at $h = 0.125$, but then falls again); the condition number of the linear system is about 3×10^4. To achieve the same error with 11 equispaced plane waves per element requires $h = 0.25$, at which point the condition number is approximately 3×10^3, a difference of one digit in accuracy for a dense linear system solve. The condition numbers of the modulated plane waves perform a little better for $k = 60$, but still exceed those of the corresponding uniform bases for the same relative error. We note that this is despite an orthogonal basis of polynomials being used for the modulated plane waves. The multiplication of the orthogonal polynomials with the plane wave part increases the overall condition numbers. Overall, in the asymptotic regime as $h \to 0$, standard plane waves perform better than a modulated plane wave basis, requiring significantly fewer degrees of freedom for a given accuracy. However, for moderate accuracy requirements (i.e. a few percent relative error), modulated plane waves can be significantly better than a standard plane wave basis in terms of degrees of freedom with similar conditioning.

6. Scattering from a unit square

In this section we will explore the application of modulated polynomials to a more difficult problem, namely sound-soft acoustic scattering of a plane-wave $e^{ikx \cdot d}$ from the unit square $S := [-0.5, 0.5]^2$. We truncate the calculation at the circle of radius 2, and approximate an absorbing boundary using an impedance boundary condition on this outer circle. The boundary conditions are $u = 0$ on the interior boundary and $\frac{\partial u}{\partial n} - iku = ik(d \cdot n - 1)e^{ikx \cdot d}$ on the outer boundary.
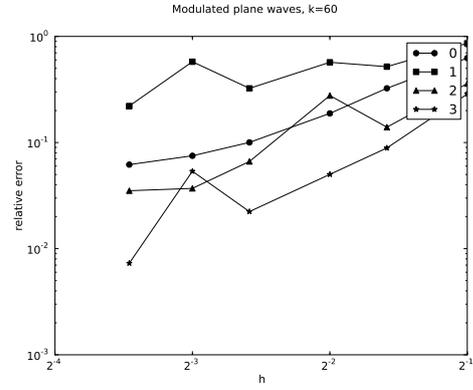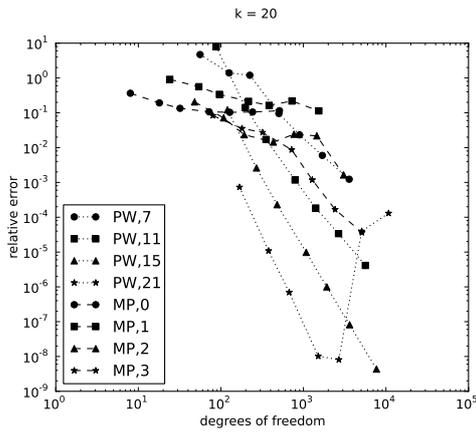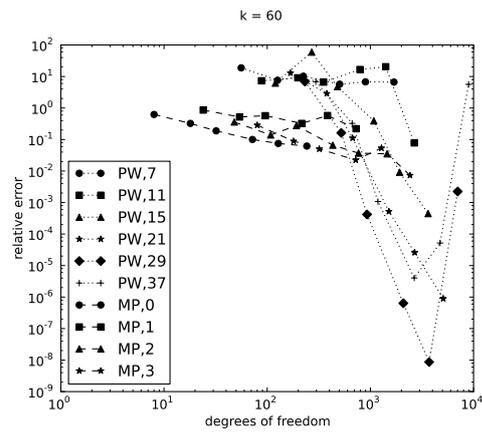
8
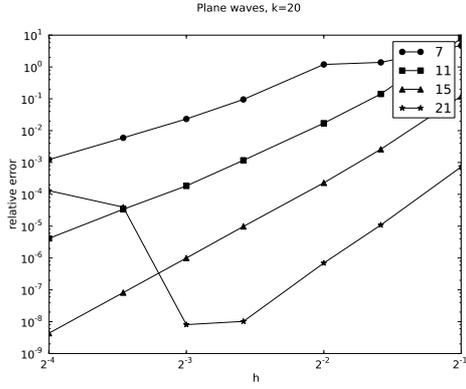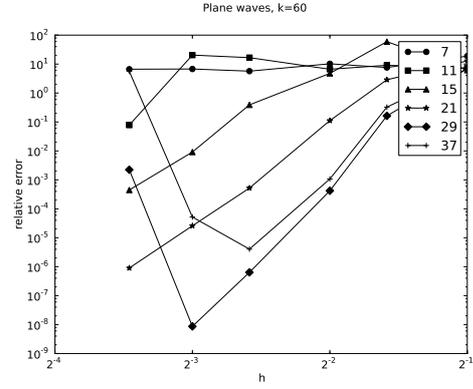
(a) Plane waves, k=20

(b) Plane waves, k=60
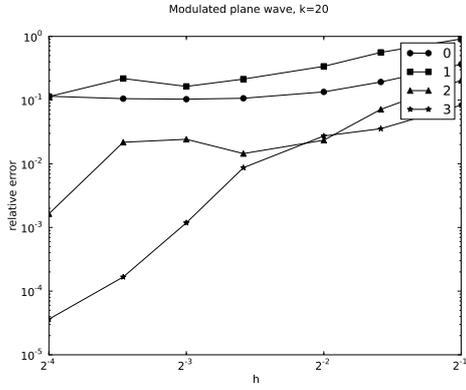
(c) Modulated waves, k=20

(d) Modulated waves, k=60

(e) relative errors vs dofs, k=20

(f) relative errors vs dofs, k=60

Figure 5: Top plots: Relative errors in the solution with plane wave basis sets for $k = 20$ (left) and $k = 60$ (right) as $h \to 0$. Middle plots: Same experiments but with a basis set consisting of modulated waves with dominant direction in each element. Bottom plots: Combination of the above error plots to show the relative error versus the number of degrees of freedom for $k = 20$ (left) and $k = 60$ (right).

9

h	No. of plane waves			
	7	11	15	21
0.5	590.7	2468	1199	394114
0.333	1600	579.6	50722	6.432×10^8
0.25	1410	3146	1726330	1.419×10^{11}
0.167	378.7	157070	4.723×10^8	4.914×10^{14}
0.125	1268	2375458	2.312×10^{10}	
0.0909	6392			

Table 1: Condition numbers for the linear system built using equispaced plane wave bases to find solutions with $k = 20$.

h	Polynomial degree			
	0	1	2	3
0.5	9.048	274.8	2488	1729
0.333	14.95	316.2	3112	8449
0.25	19.13	524.8	14414	66381
0.167	33.51	2178	32334	368706
0.125	50.53	4336	129632	391678
0.0909	76.53	11848	1321328	
0.0625	121.4	33036		

Table 2: Condition numbers for linear systems built using modulated plane wave bases to find solutions with $k = 20$.

6.1. A ray tracing based approximation

To find the dominant wave directions we use a simple ray tracing algorithm that shoots rays from the outer boundary, with initial direction d. These reflect if they hit the scatterer S. We mark elements in the mesh with the directions of the rays that intersect them. The results are shown in Figure 6.

The ray tracing creates three types of regions. Some elements are only traversed by rays associated with the incident wave. Behind the obstacle, there is a shadow region which is not traversed by any rays. The left and bottom of the square both cause reflections, which gives a diagonal band in which there are two rays, both incident and reflected.

Our first choice of basis is to take a modulated plane wave basis on each element with the directions provided by the ray tracing. In the shadow region, where there is no inherent dominant direction, we use modulated Fourier Bessel functions of the form $p(x)J_0(kr)$, where r is the polar coordinate of x local to a center in the element K. The reason for these basis functions is heuristic. Since there is no preferred wave direction from ray tracing in the shadow zone we want to use wave solutions that have an isotropic behavior in all directions, which is satisfied by the function J_0. Also, since the solution goes to zero in the shadow zone as $\omega \to \infty$ and is small compared to the other regions, a modulation by a small degree polynomial will be sufficient to give enough accuracy there.

The results, for $k = 30$, using ray tracing directions modulated with polynomials of degree 0, 2, 4 and 6 are shown in Figure 7. Since some elements now contain two modulated plane waves, ill-conditioning can occur due to these elements. To control this we remove redundant information by a local elementwise SVD. Obviously, the results are not satisfactory. The basis is missing information. The problem is that the ray traced directions do not capture corner diffraction effects as we describe in the following section.

6.2. Incorporating a GTD correction

A corner hit by a ray essentially acts like a point source, creating rays that shoot out equally in all directions. The form of this diffracted field u_d is $u_d(r) = u_i D \frac{e^{ikr}}{\sqrt{r}}$, where $u_i = A_i e^{i\psi}$ is the amplitude A_i and phase ψ_i of the incoming wave and D is a diffraction coefficient that depends on k, the direction θ of the diffracted ray and the incident phase ψ_i. A simple way of incorporating this diffraction information is to include basis functions around a corner at y_j of the form $p(x)H_0^1(k|x - y_j|)$ since $H_0^{(1)}(kr) \sim \sqrt{\frac{2}{\pi kr}}e^{i\left(kr - \frac{\pi}{4}\right)}$ as $r \to \infty$ [21]. At the corner y_j the function $H_0^{(1)}(k|x - y_j|)$ has a logarithmic singularity. However, this singularity may be removed by multiplication with a polynomial which is zero at y_j. We note that more effective basis functions around corners are possible. In [6, 22] fractional Bessel functions around corners were used to obtain an exponentially convergent basis there. But this technique was restricted to 2-d. Here, we are interested in the question whether we can obtain an acceptable for finite wavenumbers if we include a simple GTD correction term in a modulated plane wave basis.

6.3. Numerical results for the full formulation

In Figure 8 we plot the real part of the solution for $k = 30$ using the ray traced modulated plane wave basis, in the illuminated area, modulated Bessel functions of order zero in the shadow zone and modulated Hankel functions of order zero around the corners. Each modulated Hankel function is active in those elements that are in the line of
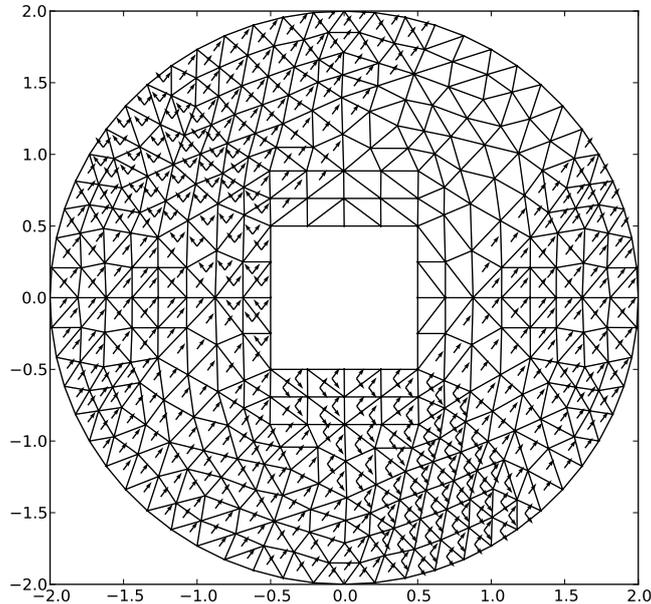
Figure 6: Ray traced directions for an incident plane wave with direction $d = [1, \ 1]^T$.

sight of the associated corner. For comparison in Figure 9 we show the solution of the same problem with 15 equally spaced plane wave directions in each element. In Figure 10 we plot the relative pointwise errors of the computed solution against a reference solution computed with 27 equally spaced plane waves in each element for the cases of a) the full modulated basis including GTD correction terms, b) a plane wave basis with 15 uniformly spaced plane waves and c) a polynomial basis with polynomials of degree 6 in each element. The reference solution is shown in plot d). It is visible that the plane wave basis has the best error behaviour. The modulated basis has large relative errors in the shadow zone. However, the reference solution is small in shadow. The degree 6 polynomials seem to behave similarly to the modulated basis set. The relative L^2 errors compared to the reference solutions are as follows: a) 3.98%, b) 0.69%, c) 4.46%. In comparison for the number of degrees of freedom we have a) 19352, b) 8760, c) 16352. Hence, the accuracy of the modulated wave basis multiplied with polynomials of degree 3 is comparable to a full polynomial basis of degree 6, but needs slightly more degrees of freedom. The plane wave basis outperforms both basis sets significantly in terms of error and required degrees of freedom. The plane wave basis has 15 degrees of freedom on each element. The modulated ray traced plane waves require 10 dofs per element in the case of one dominant direction and 20 dofs in the case of two dominant directions. Additionally, every element is within the line of sight of either 2 or 3 corners, and the modulated Hankel function for each of these contributes an additional 10 dofs per element.

In very simple situations, such as the one in Section 5, a modulated basis with dominant directions can outperform a plane wave basis. For more complicated situations, such as that presented in this example, plane waves equally spaced in all directions are preferable since tracking all dominant wave directions including GTD effects can quickly become expensive. The situation could be improved by using the exact correction term from the asymptotic theory of diffraction and / or being more sophisticated about the choice of polynomial degree used to modulate these terms. However, the analytic information required to do this may not be available for every possible configuration.
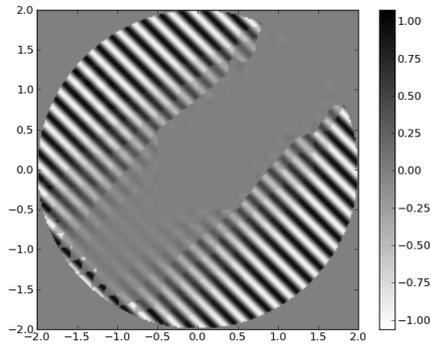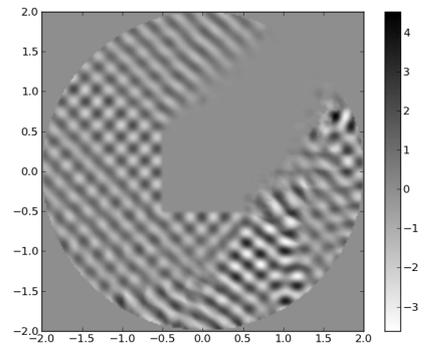
(a) p=0
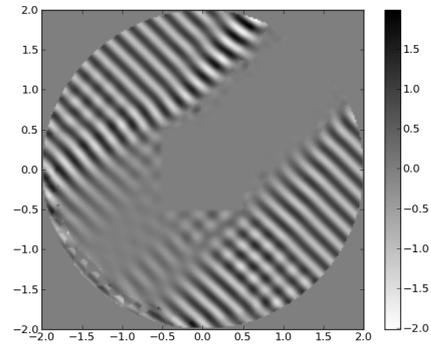
(b) p=2
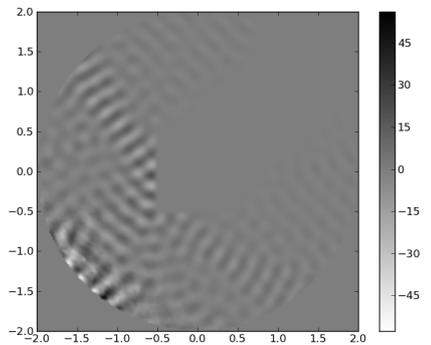
(c) p=4

(d) p=6

Figure 7: Real part of scattering solution with ray traced plane waves modulated with polynomials of degree 0, 2, 4 and 6.
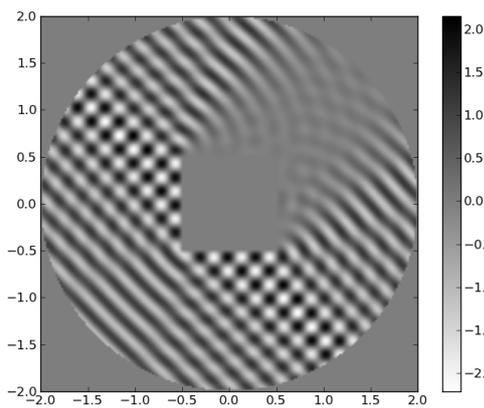

Figure 8: Real part of scattering solution with ray traced plane waves and Fourier-Hankels modulated with polynomials of degree 3
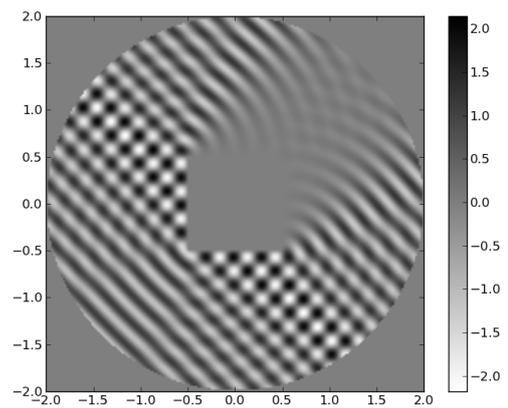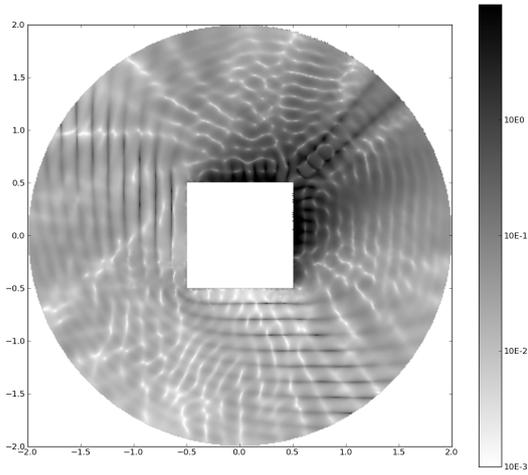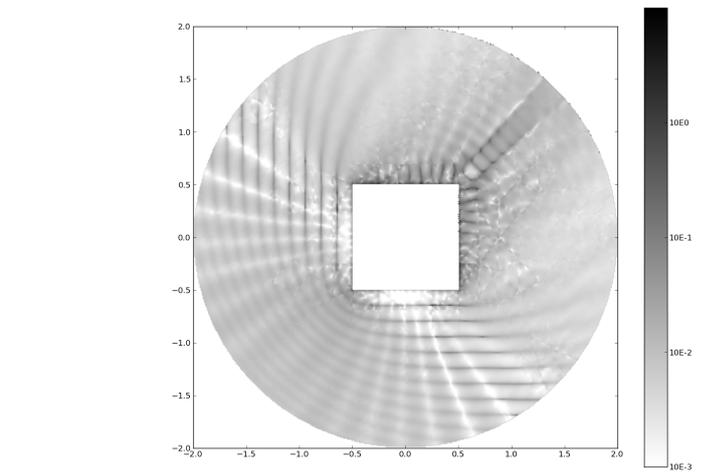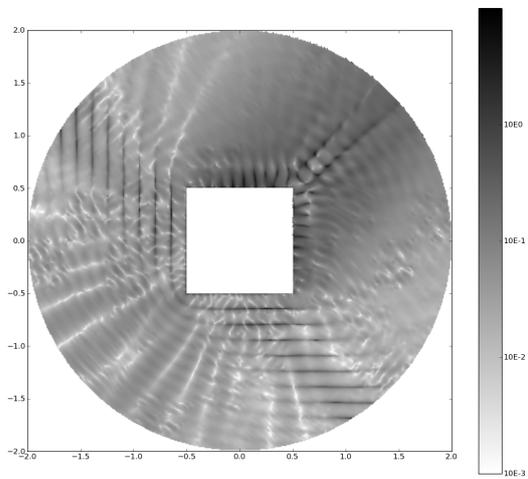

Figure 9: Real part of scattering solution with 15 uniformly spaced plane-waves per element
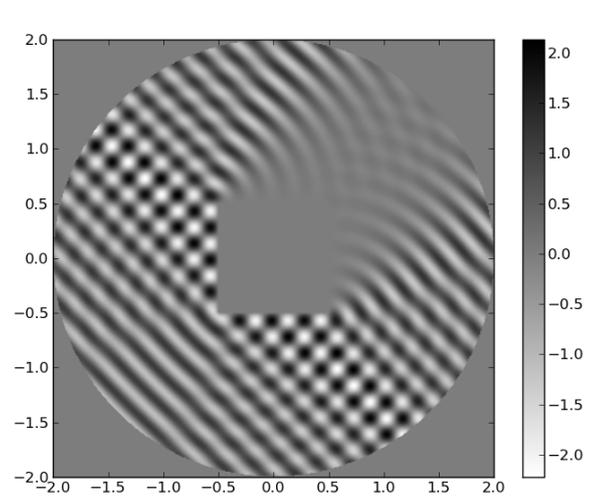
(a) Modulated basis

(b) 15 plane waves

(c) Degree 6 polynomials

(d) Reference Solution

Figure 10: Pointwise relative errors for a) waves modulated with polynomials of degree 3, b) 15 equally spaced plane waves in each element and c) a polynomial basis of degree 6 polynomials in each element. For comparison the reference solution is shown in d).
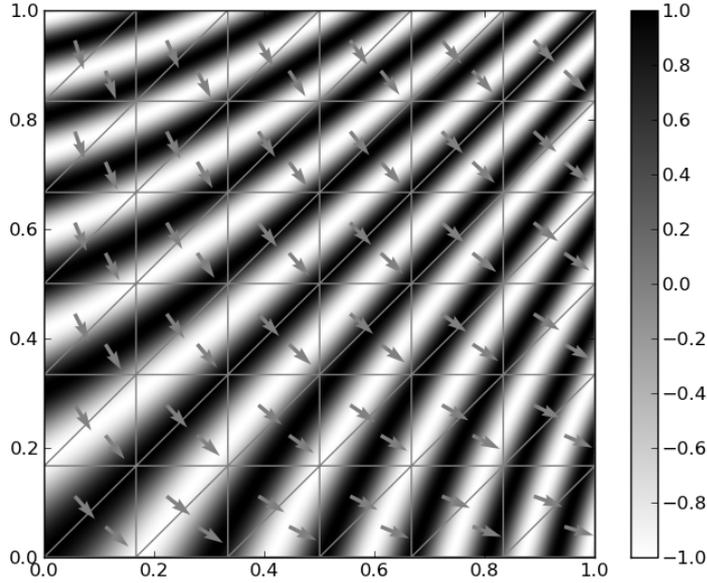
Figure 11: Real part of the solution to the inhomogeneous problem and dominant wave directions in each element.

7. The case of varying speed of sound

In this section we investigate an example of a problem with varying speed of sound. Plane wave bases only give higher order convergence for problems with constant or piecewise constant speed of sound c. For functions that are not Helmholtz solutions with constant c their approximation quality is in general no better than that of piecewise linear basis functions as $h \to 0$ [12, Section 3.2]. In this section we demonstrate that modulated dominant directions give higher order convergence and can significantly outperform both standard plane waves and polynomial bases for problems with varying c.

The idea is, similarly to the case with constant speed of sound, to approximate a Helmholtz solution of the form $u = \tilde{A}(x)e^{i\omega\phi(x)}$ in an element K by a modulated plane wave of the form $p(x)e^{i\frac{\omega}{\hat{c}}d \cdot x}$, where in the case of varying $c(x)$ we choose $\hat{c} = c(x_0)$ for some $x_0 \in K$. For problems with moderately changing c this provides a sufficient phase approximation across the whole element.

If ψ is a harmonic function, then a simple calculation shows that the function $u(x) = e^{i\omega\psi(x)}$ satisfies

$$\Delta u + \omega^2 |\nabla\psi|^2 u = 0.$$

We choose $\psi(x) = \frac{3}{2\sqrt{2}} \left[(x + \frac{1}{3})^2 - (y + \frac{1}{3})^2 \right]$. Hence, we have

$$c(x) = \sqrt{\frac{2}{(3x + 1)^2 + (3y + 1)^2}}$$

The parameters are chosen in such a way that $c(x)$ varies between $\frac{1}{4}$ and 1 in the square $[0, 1]^2$. On the boundary, we specify Dirichlet conditions compatible with the analytic solution $u = e^{i\omega\phi(x)}$. We fix $\omega = 20$ and discretise using a regular triangular mesh.

Since we know the analytic solution we can compute the exact gradient $\nabla\psi(x)$ in the center point of each element to obtain the dominant direction. This is shown in Figure 11. Figure 12 shows the rate of convergence for decreasing h in the case of a modulated plane wave basis with degree 2 polynomials (circles), a standard polynomial basis of degree 2
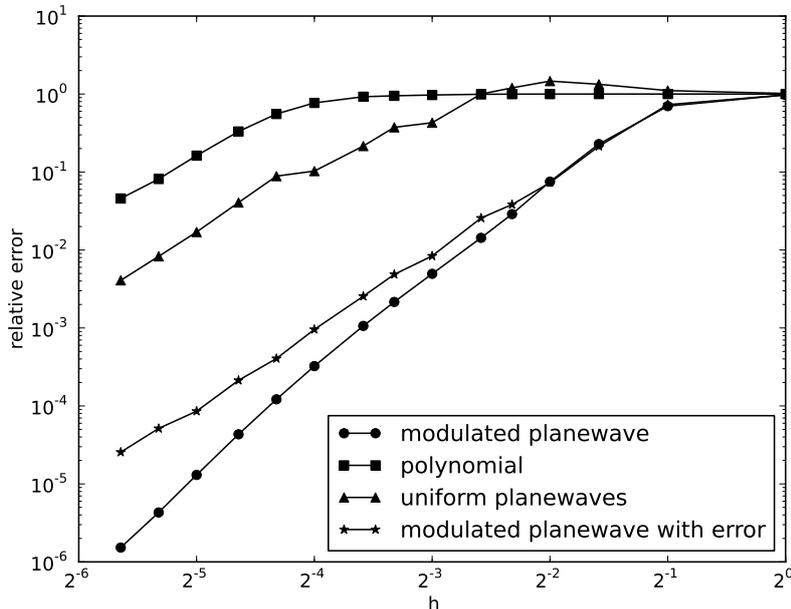
14

Figure 12: *h*-convergence of various bases for the inhomogeneous problem.

(squares), 15 uniform plane wave directions in each element (triangles) and a modulated plane wave basis with degree 2 polynomials, where the dominant directions have a random error of 1% (stars). It is obvious that modulated plane waves perform significantly better than the other basis types. Already for $h = 2^{-3}$, which corresponds to around two wavelengths at the upper right part of the unit square, we have a relative L^2 error of less than 1%, using only 6 dofs in each element. Perturbing the optimal directions only increases the error slightly. For comparison, the error of the plane wave basis is around two orders of magnitude larger than the error of the modulated basis for $h = 2^{-3}$.

8. Conclusions

In this paper we asked how well we can approximate the solutions of Helmholtz problems using just plane waves in a dominant direction instead of the approximately equally distributed plane waves that are typically used in plane wave methods. In order to achieve higher order convergence we need to modulate the dominant directions with polynomials. In 2-d a polynomial basis of degree p needs $O(p^2)$ basis functions, while in 3-d already $O(p^3)$ basis functions are required. For Helmholtz problems with constant c equally distributed plane waves only need $O(p)$ basis functions in 2-d and $O(p^2)$ basis functions in 3-d to achieve the same asymptotic approximation quality as polynomials of degree p as $h \to 0$. Hence, although increasing the degree of the polynomials in modulated plane wave bases does improve the order of convergence, for problems with constant speed of sound they can only be effective compared to plane waves if there are are only few dominant directions in each element and the degree of the polynomial modulation stays small. One such example was given in Section 5. For more complicated problems with varying wave directions and diffracted wave components, as was the case for the square scatterer in Section 6, modulated plane waves using dominant directions are outperformed by standard plane wave bases.

The situation changes significantly for problems with varying speed of sound. Here, plane waves do not give higher order convergence. But modulated plane wave bases continue to do so. Effectively, the modulation corrects for the varying wavenumber. The authors are currently investigating the application of modulated plane waves to more complicated problems in varying media based on dominant directions obtained from ray tracing.

15

References

[1] O. Cessenat, B. Despres, Application of an Ultra Weak Variational Formulation of Elliptic PDEs to the Two-Dimensional Helmholtz Problem, SIAM J. Numer. Anal. 35 (1) (1998) 255–299.

[2] T. Huttunen, P. Monk, J. P. Kaipio, Computational Aspects of the Ultra-Weak Variational Formulation, J. Comp. Phys. 182 (1) (2002) 27–46.

[3] G. Gabard, Discontinuous Galerkin methods with plane waves for time-harmonic problems, J. Comp. Phys. 225 (2) (2007) 1961–1984.

[4] O. Laghrouche, P. Bettess, R. J. Astley, Modelling of short wave diffraction problems using approximating systems of plane waves, Intern. J. Numer. Methods Eng. 54 (10) (2002) 1501–1533.

[5] A. Moiola, R. Hiptmair, I. Perugia, Plane wave approximation of homogeneous Helmholtz solutions, Z. Angew. Math. Phys. 62 (5) (2011) 809–837.

[6] M. Stojek, Least-squares Trefftz-type elements for the Helmholtz equation, Intern. J. Numer. Methods Eng. 41 (5) (1998) 831–849.

[7] P. Monk, D.-Q. Wang, A least-squares method for the Helmholtz equation, Comput. Methods Appl. Mech. Engrg. 175 (1-2) (1999) 121–136.

[8] A. H. Barnett, T. Betcke, Quantum mushroom billiards, Chaos 17 (043125).

[9] G. Fairweather, A. Karageorghis, The method of fundamental solutions for elliptic boundary value problems, Adv. Comput. Math. 9 (1-2) (1998) 69–95.

[10] A. H. Barnett, T. Betcke, Stability and convergence of the method of fundamental solutions for Helmholtz problems on analytic domains, J. Comp. Phys. 227 (7003–7026).

[11] E. Giladi, J. B. Keller, A Hybrid Numerical Asymptotic Method for Scattering Problems, J. Comp. Phys. 174 (1) (2001) 226–247.

[12] C. J. Gittelson, R. Hiptmair, I. Perugia, Plane wave discontinuous Galerkin methods: analysis of the h-version, M2AN Math. Model. Numer. Anal. 43 (2) (2009) 297–331.

[13] B. Engquist, O. Runborg, Computational high frequency wave propagation, Acta Numer. 12 (-1) (2003) 181–266.

[14] V. Vinje, E. Iversen, H. var Gjø ystdal*, Traveltime and amplitude estimation using wavefront construction, Geophysics 58 (8) (1993) 1157–1166.

[15] S. Kim, C.-S. Shin, J. B. Keller, High-frequency asymptotics for the numerical solution of the Helmholtz equation, Appl. Math. Lett. 18 (7) (2005) 797–804.

[16] J. M. Melenk, I. Babuska, The partition of unity finite element method: Basic theory and applications, Comput. Methods Appl. Mech. Engrg. 139 (1-4) (1996) 289–314.

[17] I. Babuska, J. M. Melenk, The partition of unity method, Intern. J. Numer. Methods Eng. 40 (4) (1997) 727–758.

[18] C. Farhat, R. Tezaur, J. Toivanen, A domain decomposition method for discontinuous Galerkin discretizations of Helmholtz problems with plane waves and Lagrange multipliers, Intern. J. Numer. Methods Eng. 78 (13) (2009) 1513–1531.

[19] D. N. Arnold, F. Brezzi, B. Cockburn, L. D. Marini, Unified Analysis of Discontinuous Galerkin Methods for Elliptic Problems, SIAM J. Numer. Anal. 39 (5) (2002) 1749–1779.

[20] S. Giani, L. Grubišić, J. S. Ovall, Benchmark results for testing adaptive finite element eigenvalue procedures, Applied Numerical Mathematics 62 (2) (2012) 121–140.

[21] S. N. Karp, A convergent 'farfield' expansion for two-dimensional radiation functions, Comm. Pure Appl. Math. 14 (1961) 427–434.

[22] A. H. Barnett, T. Betcke, An exponentially convergent nonpolynomial finite element method for time-harmonic scattering from polygons, SIAM J. Sci. Comput. 32 (3) (2010) 1417–1441.