

A phase-based hybridizable discontinuous Galerkin method for the numerical solution of the Helmholtz equation



N.C. Nguyen ^{a,*}, J. Peraire ^a, F. Reitich ^b, B. Cockburn ^b

^a Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

^b School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA

ARTICLE INFO

Article history:

Received 11 February 2014

Received in revised form 29 January 2015

Accepted 4 February 2015

Available online 10 February 2015

Keywords:

Finite element method
Discontinuous Galerkin methods
Hybrid/mixed methods
Geometrical optics
Helmholtz equation
Eikonal equation

ABSTRACT

We introduce a new hybridizable discontinuous Galerkin (HDG) method for the numerical solution of the Helmholtz equation over a wide range of wave frequencies. Our approach combines the HDG methodology with geometrical optics in a fashion that allows us to take advantage of the strengths of these two methodologies. The phase-based HDG method is devised as follows. First, we enrich the local approximation spaces with precomputed phases which are solutions of the eikonal equation in geometrical optics. Second, we propose a novel scheme that combines the HDG method with ray tracing to compute multivalued solution of the eikonal equation. Third, we utilize the proper orthogonal decomposition to remove redundant modes and obtain locally orthogonal basis functions which are then used to construct the global approximation spaces of the phase-based HDG method. And fourth, we propose an appropriate choice of the stabilization parameter to guarantee stability and accuracy for the proposed method. Numerical experiments presented show that optimal orders of convergence are achieved, that the number of degrees of freedom to achieve a given accuracy is independent of the wave number, and that the number of unknowns required to achieve a given accuracy with the proposed method is orders of magnitude smaller than that with the standard finite element method.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

The numerical solution of wave phenomena in acoustics, elastodynamics and electromagnetics has found important applications in many areas of engineering and science such as aerospace, geophysics, civil engineering, mechanical engineering, telecommunication, medicine, and biology. Examples of applications include noise reduction, stealth technology, seismic and earthquake nondestructive testing, antenna design, the detection of hidden targets, radar, satellite, nanophotonic devices, optical fibers, waveguides, and medical imaging. The wide range of applications has led to the development of many numerical methods for simulating wave propagation phenomena over a wide frequency spectrum. Direct numerical methods such as finite element (FE), finite difference (FD), and boundary element (BE) methods are commonly used to solve problems in domains that can span tens or even a few hundred wavelengths. However, these direct numerical methods are rarely used to solve problems at high frequencies because the number of grid points required to resolve the waves with these methods increases dramatically with the wave number k .

As a consequence, asymptotic numerical methods have been developed to deal with problems at high frequencies [5]. The simplest asymptotic method is geometrical optics (GO) that expresses the solution to the wave equations as an exponential

* Corresponding author.

E-mail address: cuongng@mit.edu (N.C. Nguyen).

function times an infinite series in k^{-1} . This asymptotic expansion is substituted into the wave equations to yield the eikonal equation for the phase and a sequence of transport equations for the amplitudes [48]. The main drawbacks of the GO model are that it does not account for diffraction effects at boundaries and breaks down at caustics where the predicted amplitude is unbounded. More sophisticated models such as the geometrical theory of diffraction (GTD) [5] and uniform theory of diffraction (UTD) [35] are needed to bring diffraction phenomena into the GO model. The main advantage of asymptotic numerical methods is that their complexity is independent of the wave number k , while their accuracy generally improves as k increases. However, because these methods approximate the wave equations at infinite frequency, they tend to be less accurate for moderate frequencies and may lead to erroneous results for problems at the low frequency spectrum.

In this paper, we introduce a new hybridizable discontinuous Galerkin (HDG) method for the numerical solution of the Helmholtz equation over a wide range of wave frequencies. Our approach combines the HDG methodology [13,19,30,42,43] with geometrical optics in a fashion that allows us to take advantage of the strengths of these two methodologies. The new HDG method is devised as follows. First, we enrich the local approximation spaces with precomputed phases which are solutions of the eikonal equation in geometrical optics. Second, we devise a systematic procedure for computing multivalued solutions of the eikonal equation. Third, we utilize the proper orthogonal decomposition to remove redundant modes and obtain locally orthogonal basis functions which are then used to construct the global approximation spaces of the HDG methods. These basis functions allow for a good representation of the solution of the Helmholtz equation because they are able to capture the oscillatory part of the solution for a wide range of wave numbers. And fourth, we propose an appropriate choice of the stabilization parameter to guarantee stability and accuracy. We present several numerical examples to demonstrate the performance of the proposed approach. Numerical results show that optimal orders of convergence are achieved and that the number of unknowns required to achieve a given accuracy with the proposed method is orders of magnitude less than that with the standard finite element method.

The idea of enriching the approximation spaces with wave-like functions has been applied to both FE methods and BE methods. One approach is that a large number of free-space wave solutions are used to construct the basis functions. For instance, the basis functions are the products of polynomials and plane waves which are uniformly distributed on the unit circle (in 2D) or unit sphere (in 3D). This is the approach pursued by partition of unity [2], ultra weak variational formulation [8], discontinuous enrichment methods [25], and plane wave DG method [28,32]. Although these methods can reduce the number of grid points necessary to resolve the waves, they suffer from inherent ill-conditioning due to the choice of the basis functions. Their advantage compared to standard FE methods is not clear. Another hybrid approach constructs the basis functions as the products of polynomials and the oscillatory phase factors, just as we do, instead of plane waves. This approach has been applied in both the BE context [6,34] and in the FE context [29]. Because its basis functions capture most of the oscillatory part of the solution, it needs significantly less grid points than the previous approach to achieve the same accuracy. Like the methods proposed here, they require the computation of multiple solutions of the eikonal equation — a challenging problem — that has been extensively investigated by several researchers and solved by using Lagrangian ray tracing schemes [1,4] or Eulerian schemes [18,24,26,44]. In contrast with the method proposed here, the methods developed in [6,34] are based on integral equations and, as such, are limited e.g. to piecewise constant material properties (precluding, for instance, simulations in hydroacoustics). Moreover, our variational formulation avoids the need for singular integration, which may be further complicated by highly oscillatory behavior. On the other hand, the formulation of the method [29] differs with ours only in the type of finite element method: Instead of the continuous Galerkin method, we use the HDG method. However, our contribution is unique in a number of aspects including that it is the first successful application to smooth non-convex configurations giving rise to multiple scattering; that it uses HDG with all its advantages; and that it uses HDG to ascertain the GO solution.

The paper is organized as follows. In Section 2, we present a short overview of the literature to put the contribution of this paper in perspective. In Section 3, we introduce our phase-based HDG method. In Section 4, we present several numerical examples to demonstrate the performance of the proposed approach. Finally, in Section 5, we end the paper with some concluding remarks.

2. Literature overview

2.1. HDG methods

Although the DG methods were proven to be successful for a variety of nonlinear hyperbolic problems [21], their straightforward application to diffusion problems was criticized for providing sub-optimally convergent approximations for the flux, in contrast with mixed methods, as well as for producing a substantially larger amount of globally-coupled degrees of freedom (for the same mesh and polynomial degree of the approximation) in comparison to the standard continuous Galerkin method. The HDG methods [13] were introduced to address those criticisms. Indeed, these methods were devised to guarantee that only the degrees of freedom of the approximation of the scalar variable on the interelement boundaries are globally coupled. Later, they were proven to provide an optimal order of convergence for the approximate flux and to share with the mixed methods their superconvergence properties for the scalar variable [11,15,16]. In this manner, the HDG methods retain the advantages typical of DG methods, namely, the ease in handling variable-degree approximations and nonconforming meshes, while being as efficiently implementable and accurate as the best mixed methods. This, in turn, motivated and fueled the extension of the HDG methods to a variety of steady-state and time-dependent problems of different types:

diffusion problems [10,33], convection–diffusion problems [12,38,39], incompressible flow [14,17,20,40,41], continuum mechanics [7,36,37,45,50], and, more recently, to time-dependent acoustic and elastic wave propagation in [19,42] and to the time-harmonic Maxwell's equations [43] and the Helmholtz equation in [27,30]. In the setting of wave propagation problems, the HDG methods compare with other finite element methods favorably because they achieve optimal orders of convergence for both the scalar and vector unknowns and display new superconvergence properties.

We briefly describe HDG methods for a model Helmholtz equation

$$-\Delta u - k^2 u = f \quad \text{in } \Omega \subset \mathbb{R}^d, \quad (1a)$$

$$\partial_n u + ik\alpha u = g \quad \text{on } \partial\Omega, \quad (1b)$$

where Ω is a Lipschitz domain in \mathbb{R}^d ($d \geq 2$). In (1b), the coefficient α varies on the boundary $\partial\Omega$ and represents different types of boundary conditions. Specifically, the Neumann boundary condition corresponds to $\alpha = 0$, the Dirichlet boundary condition to $\alpha = \infty$, and the first-order absorbing boundary condition to $\alpha = \pm 1$. The first-order absorbing boundary condition in (1b) is taken solely for the purposes of illustration, as it can be readily replaced by higher-order local or exact global (Dirichlet-to-Neumann) conditions or, more simply, by suitable perfectly matched layers (PML).

We next rewrite the Helmholtz equation as a first-order system

$$\mathbf{q} - \nabla u = 0, \quad \text{in } \Omega, \quad (2a)$$

$$-\nabla \cdot \mathbf{q} - k^2 u = f, \quad \text{in } \Omega, \quad (2b)$$

$$\mathbf{q} \cdot \mathbf{n} + ik\alpha u = g \quad \text{on } \partial\Omega. \quad (2c)$$

The physical domain Ω is triangulated into elements K forming a mesh \mathcal{T}_h satisfying the standard finite element conditions [9] (see p. 38). Then, letting $\partial\mathcal{T}_h := \{\partial K : K \in \mathcal{T}_h\}$ and denoting by \mathcal{F}_h the set of the faces F of the elements $K \in \mathcal{T}_h$, we seek a vector approximation \mathbf{q}_h to \mathbf{q} , a scalar approximation u_h to u , and a scalar approximation \hat{u}_h to the trace of u on element boundaries in spaces of the form

$$\mathbf{V}_h = \{\mathbf{v} \in \mathbf{L}^2(\Omega) : \mathbf{v}|_K \in \mathbf{V}(K) \forall K \in \mathcal{T}_h\}, \quad (3a)$$

$$W_h = \{w \in L^2(\Omega) : w|_K \in W(K) \forall K \in \mathcal{T}_h\}, \quad (3b)$$

$$M_h = \{\mu \in L^2(\mathcal{F}_h) : \mu|_F \in M(F) \forall F \in \mathcal{F}_h\}, \quad (3c)$$

respectively, where $\mathbf{V}(K)$, $W(K)$, and $M(F)$ are suitably chosen finite dimensional spaces. For instance, in the standard HDG method, we choose equal-degree local spaces

$$\mathbf{V}(K) := [P_p(K)]^d, \quad W(K) := P_p(K), \quad M(F) := P_p(F), \quad (4)$$

where $P_p(D)$ is a space of complex-valued polynomials of degree at most p on D . We introduce the following inner products

$$(v, \bar{w})_{\mathcal{T}_h} := \sum_{K \in \mathcal{T}_h} (v, \bar{w})_K, \quad \langle v, \bar{w} \rangle_{\partial\mathcal{T}_h} := \sum_{K \in \mathcal{T}_h} \langle v, \bar{w} \rangle_{\partial K}, \quad (5)$$

where we write $(u, \bar{v})_D := \int_D u \bar{v} dx$ whenever D is a domain of \mathbb{R}^d , and $\langle u, \bar{v} \rangle_D := \int_D u \bar{v} dx$ whenever D is a domain of \mathbb{R}^{d-1} . For vector-valued functions \mathbf{v} and \mathbf{w} , the integrals are similarly defined with the integrand being the dot product $\mathbf{v} \cdot \bar{\mathbf{w}}$. Note that $\bar{\mathbf{w}}$ denotes the complex conjugate of \mathbf{w} .

The HDG approximations $(\mathbf{q}_h, u_h, \hat{u}_h)$ in $W_h \times \mathbf{V}_h \times M_h$ are determined by requiring that

$$(\mathbf{q}_h, \bar{\mathbf{r}})_{\mathcal{T}_h} + (u_h, \nabla \cdot \bar{\mathbf{r}})_{\mathcal{T}_h} - (\hat{u}_h, \bar{\mathbf{r}} \cdot \mathbf{n})_{\partial\mathcal{T}_h} = 0, \quad (6a)$$

$$(\mathbf{q}_h, \nabla \bar{w})_{\mathcal{T}_h} - (\hat{u}_h, \bar{w})_{\partial\mathcal{T}_h} - (k^2 u_h, \bar{w})_{\mathcal{T}_h} = (f, \bar{w})_{\mathcal{T}_h}, \quad (6b)$$

$$\langle \hat{u}_h, \mathbf{n} + ik\alpha \hat{u}_h, \bar{\mu} \rangle_{\partial\mathcal{T}_h} = \langle g, \bar{\mu} \rangle_{\partial\Omega}, \quad (6c)$$

$$\hat{\mathbf{q}}_h = \mathbf{q}_h - \tau(u_h - \hat{u}_h)\mathbf{n} \quad \text{on } \partial\mathcal{T}_h, \quad (6d)$$

hold for all (\mathbf{r}, w, μ) in $\mathbf{V}_h \times W_h \times M_h$. Here, τ is the so-called *stabilization function*.

Note that, on any given interior face $F \in \mathcal{F}_h$, the numerical trace $\hat{\mathbf{q}}_h$ given by (6d) is double-valued since the traces on F of \mathbf{q}_h and u_h are double-valued. This is the key feature of the HDG method since this property allows us to use the first three equations (6) to express, in an elementwise manner, the functions \mathbf{q}_h , u_h and $\hat{\mathbf{q}}_h$ in terms of the data f and \hat{u}_h . The approximation \hat{u}_h is then determined by (6c), which actually enforces the single-valuedness of the normal component of the $\hat{\mathbf{q}}_h$. This is how the HDG method guarantees that only globally coupled degrees of freedom are those of the approximate trace \hat{u}_h . By means of the so-called *hybridization* technique, the HDG method (6) gives rise to the following linear system

$$\hat{\mathbf{A}}\hat{\mathbf{u}} = \mathbf{f}. \quad (7)$$

Here $\hat{\mathbf{u}} \in \mathbb{C}^{N_{\text{dof}}}$ is the vector containing the degrees of freedom of \hat{u}_h with

$$N_{\text{dof}} = N_f \times N_p, \quad (8)$$

where N_f is the number of faces in the mesh and N_p is the number of polynomials per face, which is equal to $p + 1$ in two dimensions or $(p + 1)(p + 2)/2$ in three dimensions.

In [30] it was shown that, if kh is small enough and if the stabilization function τ is taken to be a uniformly bounded, strictly positive function, the approximations \mathbf{q}_h and u_h converge with the optimal order $p + 1$ whenever they are taken to be piecewise polynomials of degree $p \geq 0$ (in each component). Moreover, the computations in [30] seem to suggest that the condition number of the matrix \mathbf{A} might be independent of the wave number k , a point that seems to be corroborated by the results obtained in [27]. Indeed, in [27] two DG methods, one of them strongly related to the HDG method under consideration, were proven to be well defined independently of the value of the wave number k piecewise linear approximations. In particular, the $L^2(\Omega)$ -norm of the approximate gradient and k times the $L^2(\Omega)$ -norm of the scalar approximation were both bounded uniformly with respect to the wave number k as k goes to infinity. However, with the choice of the local spaces in (4), the global number of unknowns N_{dof} would increase significantly with the wave number k . As a result, the standard HDG method (6) is not computationally efficient for large values of k .

2.2. Asymptotic numerical methods

Asymptotic numerical methods overcome this inability of direct numerical techniques to solve the Helmholtz problem (1) at high wave numbers. One of the popular asymptotic approaches is the Luneberg–Kline expansion [5] (see p. 93) which expresses the solution of the Helmholtz problem by the following series:

$$u(\mathbf{x}) \sim e^{ik\varphi(\mathbf{x})} \sum_{n=0}^{\infty} \frac{A_n(\mathbf{x})}{k^n}, \quad (9)$$

for an unknown phase φ and unknown amplitudes A_n . Let us assume for the sake of simplicity that $f = 0$ in (1). Substituting this expression into (1a) and letting $k \rightarrow \infty$ we readily obtain the eikonal equation

$$|\nabla \varphi| = 1, \quad \text{in } \Omega, \quad (10)$$

for the phase, while the amplitudes solve a sequence of transport equations. This is the standard geometrical optics approach for numerically solving the Helmholtz equation [5].

A major drawback of the above asymptotic expansion is that it can only capture single-phase wave fields. As a result, many asymptotic methods consider a more general expansion

$$u(\mathbf{x}) \sim \sum_{n=1}^N A_n(\mathbf{x}; k) e^{ik\varphi_n(\mathbf{x})}, \quad (11)$$

where the phases φ_n are independent of the wave number k and the amplitudes A_n are mildly dependent on the wave number k . Typically, the global expansion (11) breaks down at a small set of points, namely focus points, caustics, discontinuities in the wave speed and non-smooth boundary points. Like the phase φ of the Luneberg–Kline expansion, the phases φ_n in (11) also solve the eikonal equation (10). The asymptotic expansion (11) is the point of departure for advanced asymptotic methods such as the geometrical theory of diffraction (GTD) [5] and uniform theory of diffraction (UTD) [35].

In spite of the apparent simplicity of the eikonal equation (10), its nonlinear character and the multi-valuedness of its solution, pose significant challenges to its numerical approximation and render this problem a very active area of research in computational science. Clearly a simple procedure to solve the eikonal equation can be based on the “method of characteristics” (ray-tracing) [1]. However, ray tracing has a number of issues related to ray divergence and wavefront reconstruction [24]. The consequent limitations have prompted the recent development of new computational methods based on (Eulerian) solution of partial differential equations. Early versions of this approach concentrated on the design of upwind [52,53], ENO schemes [26], and fast marching algorithms [49] for the direct solution of the eikonal equation (10). This (single-valued) solution, however, represents only the wave of first arrival at any given point and it may thus be insufficient for certain applications wherein significant effects arise as a consequence of multiple arrivals. For this reason, a number of algorithms have more recently been developed to upgrade the viscosity solution to the multi-valued solution. Among these we encounter, for instance, the big ray tracing method [4] and the slowness matching method [51]. All of these procedures are based on domain decomposition and local approximations of viscosity solutions, which are then combined into a multi-valued quantity.

On the other hand, an alternative approach to the approximation of multi-valued solutions is based on a “kinetic” formulation that views rays as trajectories of particles following a Hamiltonian dynamics [44]. In this approach, multi-valued solutions are naturally “unfolded” through the introduction of conjugate phase variables. This, however, is achieved at the expense of doubling the number of independent variables, with the consequent potential for increased computational cost. To deal with this problem, two alternative strategies have been developed, leading to “wavefront” and “moment-based” methods respectively [24]. In the former, an interface representing a wavefront is evolved following the Liouville formulation, while the latter is based on the derivation of new equations (for the moments of the density) with fewer unknowns. A recent scheme [18] sought to combine elements from these two approaches by relying on the evolution of an interface (defined in terms of level-set functions [44]) while avoiding the direct discretization of the phase variables. Instead, the

procedure is based on suitable (spectral) representations of the unknown field quantities and, as such, it can be related to moment methods where the moments are not chosen to be integrals against monomials in phase variables (needing a “closure hypothesis”), but rather against basis functions that guarantee accurate representations of general phase variations.

Once the phases have been determined by using the above techniques, one can compute the amplitudes by solving the transport equations or using GTD/UTD [5,35]. The most advantageous characteristic of these asymptotic methods is their ability to bypass the need to resolve field oscillations on the scale of the wavelength. This is, however, attained at the expense of a loss of error-controllability, which follows from the approximation that the theories entail at the level of the models (e.g., replacing the Helmholtz or Maxwell model by the eikonal equation and the transport equations). As a consequence, these high-frequency methods can incur significant inaccuracies when applied at finite frequencies, leaving a sizable gap in the range of frequencies wherein engineering or industrial design can rely on simulation [47].

2.3. Hybrid methods

Hybrid methods involve a combination of direct and asymptotic numerical methods. One simple approach is to take the local basis functions as a large number of free-space plane waves:

$$W_{\text{PW}}(K) = \text{span}\{\exp(ik\mathbf{d}_n \cdot \mathbf{x}), 1 \leq n \leq N, \forall \mathbf{x} \in K\}, \quad (12)$$

where the direction vectors $\mathbf{d}_n, 1 \leq n \leq N$, are uniformly distributed on the unit circle (in 2D) or unit sphere (in 3D). More generally, one can take the local approximation space to consists of the products of plane waves and standard finite element polynomials:

$$W_{\text{PPW}}(K) = \text{span}\{\phi_m(\mathbf{x}) \exp(ik\mathbf{d}_n \cdot \mathbf{x}), 1 \leq n \leq N, 1 \leq m \leq M, \forall \mathbf{x} \in K\}, \quad (13)$$

where $\phi_m(\mathbf{x}), 1 \leq m \leq M$, are standard finite element shape functions over the element K . This is the approach used by the partition of unity [2], ultra weak variational formulation [8], discontinuous enrichment [25], and the plane wave DG [28,32] methods. More generally, instead of using plane waves one can employ other wave functions (e.g., the Hankel functions [46]) provided that the wave functions satisfy the Helmholtz equation. The resulting methods are then called *Trefftz-based* methods [46]. Although Trefftz-based methods are more efficient than traditional FE and BE methods, the number of grid points necessary to resolve the waves with these methods depends strongly on the wave number k . This is because the Trefftz-based wave functions are free-space solutions of the Helmholtz equation which do not take the shape of the domain and the inhomogeneity of the medium into account.

A more promising hybrid approach is to take the local approximation space as:

$$W_{\text{PE}}(K) = \text{span}\{\phi_m(\mathbf{x}) \exp(ik\varphi_n(\mathbf{x})), 1 \leq n \leq N, 1 \leq m \leq M, \forall \mathbf{x} \in K\}, \quad (14)$$

where $\varphi_n(\mathbf{x}), 1 \leq n \leq N$, are the phases described in the previous subsection. This approach has been pursued in both the FE context [29] and the BE context [6,34]. Since the phases account for the shape of the domain and the inhomogeneity of the medium, this approach require significantly less grid points than the previous approach to obtain the same accurate solution at high frequencies. However, this approach requires the computation of multi-valued solutions of the eikonal equations, which can be done prior to solving the Helmholtz equation. Once the phases have been computed, they can be used to solve the Helmholtz equation for any k . The resulting methods shall be called *phase-based* methods to distinguish them from the *Trefftz-based* methods.

Last but not least, we pay attention to another hybrid approach which uses a direct numerical method for a small part of the domain (e.g., corners, cavities, shadow boundaries, caustics), and an asymptotic numerical method on the remaining part of the domain. Instances of this approach include hybrid FEM/GTD [3] and BE/GTD [23]. However, this approach requires the implementation of both the direct and the asymptotic solvers, as well as a mechanism for handling the coupling between the two solvers.

3. Phase-based HDG method

3.1. Construction of the local approximation spaces

We assume that we are given N^K phases $\varphi_n^K(\mathbf{x})$ on every element $K \in \mathcal{T}_h$ and N^F phases $\varphi_n^F(\mathbf{x})$ on every face $F \in \mathcal{E}_h$. We might attempt to define the local approximation spaces $\mathbf{V}(K)$, $W(K)$ and $M(F)$ as

$$\mathbf{V}(K) := \text{span}\{\mathbf{v}_n(\mathbf{x}) e^{ik\varphi_n^K(\mathbf{x})}, \mathbf{v}_n \in [P_p(K)]^d, n = 1, \dots, N_K\}, \quad (15a)$$

$$W(K) := \text{span}\{w_n(\mathbf{x}) e^{ik\varphi_n^K(\mathbf{x})}, w_n \in P_p(K), n = 1, \dots, N_K\}, \quad (15b)$$

$$M(F) := \text{span}\{\mu_n(\mathbf{x}) e^{ik\varphi_n^F(\mathbf{x})}, \mu_n \in P_p(F), n = 1, \dots, N_F\}. \quad (15c)$$

Note that we allow the number of phases N_K (N_F) to take different values on different elements (faces). Note also that if the phases are very similar, the basis functions of these local approximation spaces become almost linearly dependent. This in turns results in the degeneracy of the local approximation spaces and thus renders the resulting matrix system ill-conditioned and difficult to solve.

In order to avoid the degeneracy of the local approximation spaces, we orthogonalize their basis functions to obtain the corresponding orthogonalized local spaces as follows. Let us denote by $\xi_m(\mathbf{x})$, $1 \leq m \leq M_K$, the basis functions of $W(K)$, namely,

$$W(K) := \text{span}\{\xi_m(\mathbf{x}), \quad 1 \leq m \leq M_K\}. \quad (16)$$

We next introduce a symmetric matrix $\mathbf{C} \in \mathbb{C}^{M_K \times M_K}$ with entries

$$C_{ij} = \int_K \xi_i(\mathbf{x}) \bar{\xi}_j(\mathbf{x}), \quad i, j = 1, \dots, M_K. \quad (17)$$

We then consider solving the following orthonormal eigenvalue problem

$$\mathbf{C}\mathbf{v} = \lambda\mathbf{v}, \quad (18)$$

for the eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{M_K} \geq 0$ and associated eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{M_K}$. We then choose the smallest integer L_K satisfying

$$\frac{\sum_{\ell=1}^{L_K} \lambda_\ell}{\sum_{m=1}^{M_K} \lambda_m} \geq 1 - \epsilon, \quad (19)$$

where $\epsilon \geq 0$ is a specified tolerance. Finally, we compute the new basis functions

$$\zeta_\ell(\mathbf{x}) = \frac{1}{\sqrt{\lambda_\ell}} \sum_{m=1}^{M_K} v_{\ell m} \xi_m(\mathbf{x}), \quad \ell = 1, \dots, L_K, \quad (20)$$

and define an associated space

$$W_\epsilon^\perp(K) := \text{span}\{\zeta_\ell(\mathbf{x}), \quad 1 \leq \ell \leq L_K\}. \quad (21)$$

It can be easily shown that these new basis functions are orthonormal in the following sense

$$\int_K \zeta_i(\mathbf{x}) \zeta_j(\mathbf{x}) = \delta_{ij}, \quad i, j = 1, \dots, L_K. \quad (22)$$

This orthogonalization procedure is known as the proper orthogonal decomposition. Since M_K is rather small solving the eigenvalue problem (18) is inexpensive.

We note that the eigenvalue problem (18) has at least N_p positive eigenvalues, where N_p is the dimension of the space $P_p(K)$. If there is only one phase on the element K then the eigenvalue problem (18) has exactly N_p positive eigenvalues because the matrix \mathbf{C} is symmetric positive-definite. If there are multiple phases and they are nearly linearly dependent then some of the eigenvalues can be very small and close to zero. Our orthogonalization procedure rules out the use of these eigenvalues and associated eigenvectors (see Eqs. (19) and (20)). Hence, nearly linearly dependent phases do not pose any computational difficulty for the solution of the eigenvalue problem.

Similarly, we apply the POD to $V(K)$ and $M(F)$ to obtain the corresponding orthogonalized spaces $V_\epsilon^\perp(K)$ and $M_\epsilon^\perp(F)$, respectively. Note that the dimension of these orthogonalized spaces depends on ϵ . Indeed, as we reduce ϵ , we will add more orthonormal basis functions to these spaces and thus enhance their approximation properties. However, if ϵ is too small, our orthogonalized spaces may contain redundant basis functions which do not contribute to their approximation properties and yet make the resulting method more ill-conditioned. Hence, the orthogonalization serves to enhance the numerical stability of the resulting method and reduce the degrees of freedom since the redundant ones are eliminated.

3.2. Computation of multivalued solutions of the eikonal equation

For certain problems with a simple geometry, one can solve the eikonal equation *exactly* to obtain the phases. However, for most problems, we need to solve the eikonal equation numerically. As already discussed in Subsection 2.2, there are a number of different approaches for computing multiple solutions of the eikonal equation. Here we extend the approach pursued in [4] by combining ray tracing with a variation of the HDG method introduced in [39], thereby taking advantage of the strengths of both Lagrangian and Eulerian frameworks.

In order to determine the phases φ_n^K , $1 \leq n \leq N_K$, on every element $K \in \mathcal{T}_h$, we consider the problem of solving the eikonal equation with different boundary conditions:

$$|\nabla \varphi_n| = 1, \quad \text{in } K, \quad (23a)$$

$$\varphi_n = r_n, \quad \text{on } \partial K, \quad (23b)$$

for all $K \in \mathcal{T}_h$. Here the boundary data r_n , $1 \leq n \leq N_K$, to the eikonal equation are found by using the ray tracing algorithm. In practice, since the ray tracing procedure can be computationally expensive, we propose to solve the eikonal equation

at the subdomain level instead of solving it at the element level. More specifically, we first decompose the domain Ω into non-overlapping subdomains such that $\overline{\Omega} = \cup_{s=1}^S \overline{\Omega}^s$. We next compute the phases on each of the subdomains by considering the eikonal equation

$$|\nabla \varphi_n^s| = 1, \quad \text{in } \Omega^s, \quad (24a)$$

$$\varphi_n^s = r_n^s, \quad \text{on } \partial \Omega^s. \quad (24b)$$

Here the boundary data $r_n^s, 1 \leq n \leq N_s$, are obtained by the ray tracing scheme, where N_s is the number of phases on subdomain Ω^s .

Finally, we use a variation of the HDG method introduced in [39] to solve the eikonal equation (24) as follows. We rewrite the eikonal equation as a first-order system as

$$\mathbf{p}_n^s - \nabla \varphi_n^s = 0, \quad \text{in } \Omega^s, \quad (25a)$$

$$|\mathbf{p}_n^s| = 1, \quad \text{in } \Omega^s, \quad (25b)$$

$$\varphi_n^s = r_n^s, \quad \text{on } \partial \Omega^s. \quad (25c)$$

We then introduce approximation spaces

$$\mathbf{V}_h^s = \{\mathbf{v} \in \mathbf{L}^2(\Omega^s) : \mathbf{v}|_K \in [\mathcal{P}_p(K)]^d, \forall K \in \mathcal{T}_h^s\}, \quad (26a)$$

$$W_h^s = \{w \in L^2(\Omega^s) : w|_K \in \mathcal{P}_p(K), \forall K \in \mathcal{T}_h^s\}, \quad (26b)$$

$$M_h^s = \{\mu \in L^2(\mathcal{F}_h^s) : \mu|_F \in \mathcal{P}_p(F), \forall F \in \mathcal{F}_h^s\}, \quad (26c)$$

where \mathcal{T}_h^s is the finite element triangulation of Ω^s and \mathcal{F}_h^s is the collection of all faces in \mathcal{T}_h^s . We assume that $\cup_{s=1}^S \mathcal{T}_h^s = \mathcal{T}_h$.

We now seek $(\mathbf{p}_{n,h}^s, \varphi_{n,h}^s, \widehat{\varphi}_{n,h}^s) \in \mathbf{V}_h^s \times W_h^s \times M_h^s$ such that the following system of equations

$$(\mathbf{p}_{n,h}^s, \mathbf{r})_{\mathcal{T}_h^s} + (\varphi_{n,h}^s, \nabla \cdot \mathbf{r})_{\mathcal{T}_h^s} - \langle \widehat{\varphi}_{n,h}^s, \mathbf{r} \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h^s} = 0, \quad (27a)$$

$$(|\mathbf{p}_{n,h}^s|, w)_{\mathcal{T}_h^s} + \langle (\varphi_{n,h}^s - \widehat{\varphi}_{n,h}^s), w \rangle_{\partial \mathcal{T}_h^s} = (1, w)_{\mathcal{T}_h^s}, \quad (27b)$$

$$\langle (\varphi_{n,h}^s - \widehat{\varphi}_{n,h}^s), \mu \rangle_{\partial \mathcal{T}_h^s \setminus \partial \Omega^s} + \langle \widehat{\varphi}_{n,h}^s - r_n^s, \mu \rangle_{\partial \Omega^s} = 0, \quad (27c)$$

hold for all $(\mathbf{r}, w, \mu) \in \mathbf{V}_h^s \times W_h^s \times M_h^s$. In Eq. (27b) the second integral represents the jump across the element interface and plays the role of a stabilization mechanism in the HDG method. In Eq. (27c) the first integral weakly enforces the continuity of this jump across the element interface, while the second integral weakly imposes the Dirichlet boundary condition. The nonlinear system (27) can be solved by using the Newton method. At every Newton iteration, we can locally eliminate the degrees of freedom of $(\mathbf{p}_{n,h}^s, \varphi_{n,h}^s)$ to arrive at a global matrix system involving the degrees of freedom of $\widehat{\varphi}_{n,h}$ only; see [39] for a detailed discussion.

Once $(\varphi_{n,h}^s, \widehat{\varphi}_{n,h}^s)$ has been computed, we obtain φ_n^K and φ_n^F as follows. For any element $K \in \mathcal{T}_h$, we identify a subdomain Ω^ℓ containing K and set

$$\varphi_n^K(\mathbf{x}) = \varphi_{n,h}^\ell(\mathbf{x})|_K, \quad n = 1, \dots, N_\ell. \quad (28)$$

Similarly, for any face $F \in \mathcal{F}_h$, we identify a skeleton \mathcal{F}_h^ℓ containing F and set

$$\varphi_n^F(\mathbf{x}) = \widehat{\varphi}_{n,h}^\ell(\mathbf{x})|_F, \quad n = 1, \dots, N_\ell. \quad (29)$$

With these phases we construct the locally orthogonal approximation spaces as described in the previous subsection.

3.3. Phase-based HDG approximation of the Helmholtz equation

We are ready to describe our phase-based HDG approximation of the Helmholtz equation. We find $(\mathbf{q}_h, u_h, \widehat{u}_h) \in \mathbf{V}_{h,\epsilon}^\perp \times W_{h,\epsilon}^\perp \times M_{h,\epsilon}^\perp$ such that the following equations

$$(\mathbf{q}_h, \bar{\mathbf{r}})_{\mathcal{T}_h} + (u_h, \nabla \cdot \bar{\mathbf{r}})_{\mathcal{T}_h} - \langle \widehat{u}_h, \bar{\mathbf{r}} \cdot \mathbf{n} \rangle_{\partial \mathcal{T}_h} = 0, \quad (30a)$$

$$(\mathbf{q}_h, \nabla \bar{w})_{\mathcal{T}_h} - \langle \widehat{\mathbf{q}}_h \cdot \mathbf{n}, \bar{w} \rangle_{\partial \mathcal{T}_h} - (k^2 u_h, \bar{w})_{\mathcal{T}_h} = (f, \bar{w})_{\mathcal{T}_h}, \quad (30b)$$

$$\langle \widehat{\mathbf{q}}_h \cdot \mathbf{n} + ik\alpha \widehat{u}_h, \bar{\mu} \rangle_{\partial \mathcal{T}_h} = \langle g, \bar{\mu} \rangle_{\partial \Omega}, \quad (30c)$$

$$\widehat{\mathbf{q}}_h = \mathbf{q}_h - \tau (u_h - \widehat{u}_h) \mathbf{n} \quad \text{on } \partial \mathcal{T}_h, \quad (30d)$$

hold for all $(\mathbf{r}, w, \mu) \in \mathbf{V}_{h,\epsilon}^\perp \times W_{h,\epsilon}^\perp \times M_{h,\epsilon}^\perp$. Here the global approximation spaces are defined as

$$\mathbf{V}_{h,\epsilon}^\perp = \{\mathbf{v} \in \mathbf{L}^2(\Omega) : \mathbf{v}|_K \in \mathbf{V}_\epsilon^\perp(K), \forall K \in \mathcal{T}_h\}, \quad (31a)$$

$$W_{h,\epsilon}^\perp = \{w \in L^2(\Omega) : w|_K \in W_\epsilon^\perp(K), \forall K \in \mathcal{T}_h\}, \quad (31b)$$

$$M_{h,\epsilon}^\perp = \{\mu \in L^2(\mathcal{F}_h) : \mu|_F \in M_\epsilon^\perp(F), \forall F \in \mathcal{F}_h\}. \quad (31c)$$

The weak form of the phase-based HDG method (30) is similar to that of the standard HDG method (6) except for the approximation spaces. This difference is key because, as we are going to see, it is the main reason why the phase-based HDG method needs significantly less number of elements than the standard HDG method for the same accuracy. This will result in a considerable reduction in the CPU time and memory storage with the phase-based HDG method.

Finally, we note that the choice of the stabilization parameter is crucial to ensure the stability and accuracy of the phase-based HDG method. We propose to choose the stabilization parameter as

$$\tau = \pm ik, \quad (32)$$

where the sign of τ takes the sign of the coefficient α in the Robin boundary condition (1b). This choice is based on the following dimensionality argument. We note from (30d) that the dimensionality of τ must be $[1/L]$, where L is a length unit. Hence, the formula (32) guarantees the correct dimensionality of the numerical flux $\hat{\mathbf{q}}_h$. Moreover, the stabilization parameter (32) is complex and has the same sign as α because we would like it to act as damping to provide stability for the resulting scheme.

Note that in [27] (for piecewise linear approximations and for their second method with $\beta = k/2$ and $\delta = 1/(2k)$) and in [22] (for piecewise polynomials of arbitrary degree), it has been shown that our HDG method is well defined independently of the value of the wave number k in the case $\alpha = +1$ and $\tau = +ik$. Since here we are using $\alpha = -1$, this means that the method is well defined provided we take $\tau = -ik$. It is reasonable to surmise that the same result will hold for the choice of basis we are proposing in this paper. Moreover, it has been proven [30] (see also [22]) that the HDG method using piecewise polynomial approximations is optimally convergent and superconvergent for $\tau = +ik$ (and $\alpha = 1$) whenever hk is small. However, none of the above-mentioned error estimates show that the resulting method is free of the so-called pollution effect. Although the boundedness of the present method and its error estimates are a hard, open problem, we show here that, with our special choice of bases, this phenomenon can be significantly reduced.

4. Numerical experiments

In this section, we present a variety of numerical examples to demonstrate the performance of the proposed method. The first example is devised to study the convergence rate of the method. The other two examples involve sound-hard scattering from a circular cylinder and kite-shaped scatterer. They serve to demonstrate the ability of the proposed method for dealing with diffraction at convex and concave boundaries, creeping rays, and caustics. In these examples, we will compare the phase-based HDG method with the standard finite element method [9] in terms of the global number of unknowns and accuracy.

4.1. A test case

We consider solving the Helmholtz equation (1a) in a domain $\Omega = (-1/2, 1/2) \times (-1/2, 1/2)$ with a Dirichlet boundary condition $u = g_D$ on $\partial\Omega$. Here we choose the source term f and boundary data g_D such that the problem has the following exact solution

$$u_{\text{ex}} = A(x, y) \exp(ik\varphi(x, y)),$$

with $A(x, y) = \exp(x) + \exp(y)$ and $\varphi(x, y) = \sqrt{2}(x + y)/2$. Through this test case we will study the convergence of the phase-based HDG method. For this purpose, we assume that we are given the exact phase $\varphi(x, y)$.

We consider triangular meshes obtained by splitting a regular $n \times n$ Cartesian grid into a total of $2n^2$ triangles, giving uniform element sizes of $h = 1/n$. Furthermore, we use polynomials of degree p to define the local approximation spaces in (15) for all the elements of those meshes. We present in Table 1 the relative error and the order of convergence of the approximate solution u_h for $k = 10^1, 10^2, 10^3$, and 10^4 . These results are obtained with the stabilization parameter $\tau = -ik$. We observe that the convergence rate of the approximate solution is optimal with order $O(h^{k+1})$ in the L^2 -norm for all of these values of k . We emphasize that these convergence results are obtained for very coarse meshes and high wavenumbers.

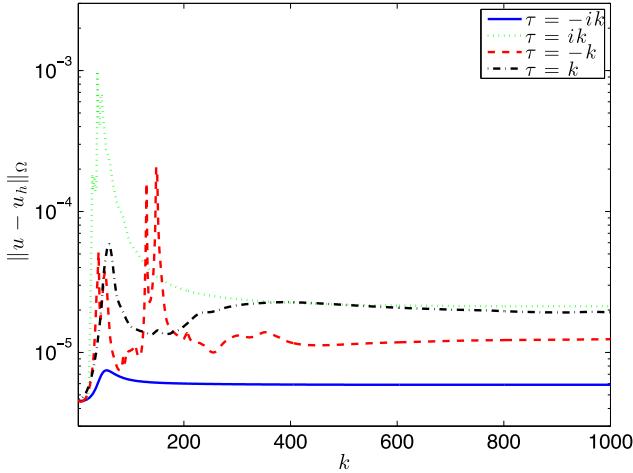
To understand the effect of the stabilization on the accuracy, we consider solving the same problem with the following Robin boundary condition (instead of the Dirichlet boundary condition)

$$\nabla u \cdot \mathbf{n} + ik\alpha u = g_R, \quad \text{on } \partial\Omega, \quad (33)$$

for $\alpha = -1$, where we choose g_R so as to have the same exact solution as given above. We display in Fig. 1 the L^2 error as a function of k for four different choices of τ when we take $h = 0.1$ and $p = 2$. We observe that the best results are obtained with $\tau = -ik$, whereas the worst results are obtained with $\tau = +ik$. Furthermore, the choice $\tau = -ik$ gives very similar results for both the Robin boundary condition and the Dirichlet boundary condition. These results justify our choice of the stabilization parameter according to (32). We shall thus use it in subsequent numerical experiments.

Table 1Relative error and order of convergence of the approximate solution u_h for $\tau = -ik$.

Degree p	Mesh h	$k = 10$		$k = 100$		$k = 1000$		$k = 10\,000$	
		Error	Order	Error	Order	Error	Order	Error	Order
1	1/2	1.79e-2	–	1.94e-2	–	1.93e-2	–	1.93e-2	–
	1/4	3.94e-3	2.19	4.80e-3	2.01	4.76e-3	2.02	4.76e-3	2.02
	1/8	9.14e-4	2.11	1.18e-3	2.02	1.18e-3	2.01	1.18e-3	2.01
	1/16	2.22e-4	2.04	2.80e-4	2.08	2.94e-4	2.01	2.94e-4	2.01
	1/32	5.50e-5	2.01	6.20e-5	2.17	7.34e-5	2.00	7.32e-5	2.00
2	1/2	8.89e-4	–	7.94e-4	–	7.73e-4	–	7.73e-4	–
	1/4	7.60e-5	3.55	9.98e-5	2.99	9.51e-5	3.02	9.51e-5	3.02
	1/8	8.90e-6	3.10	1.30e-5	2.94	1.18e-5	3.01	1.18e-5	3.01
	1/16	1.10e-6	3.01	1.47e-6	3.14	1.47e-6	3.00	1.47e-6	3.01
	1/32	1.37e-7	3.00	1.57e-7	3.23	1.87e-7	2.98	1.82e-7	3.01
3	1/2	3.89e-5	–	2.44e-5	–	2.34e-5	–	2.34e-5	–
	1/4	1.12e-6	5.12	1.49e-6	4.03	1.45e-6	4.02	1.44e-6	4.02
	1/8	6.77e-8	4.04	9.00e-8	4.05	9.03e-8	4.00	8.91e-8	4.01
	1/16	4.21e-9	4.01	5.07e-9	4.15	5.62e-9	4.01	5.55e-9	4.00
	1/32	2.63e-10	4.00	2.86e-10	4.15	3.51e-10	4.00	3.48e-10	4.00

**Fig. 1.** The L^2 error as a function of k for four different choices of τ with $h = 0.1$ and $p = 2$. This result is obtained with the Robin boundary condition (33) instead of the Dirichlet boundary condition.

4.2. Sound-hard scattering from a circular cylinder

We consider the exterior Helmholtz problem for scattering of an incident plane wave $u^{\text{inc}} = \exp(ikx)$ by a circle of radius $a = 1$. The geometry and boundary conditions are depicted in Fig. 2(a), while the finite element mesh used in the computation is shown in Fig. 2(b). We impose the Neumann condition $\nabla u \cdot \mathbf{n} = -\nabla u^{\text{inc}} \cdot \mathbf{n}$ on the unit circle and the Robin condition $\nabla u \cdot \mathbf{n} - iku = \mathbf{g}$ on the outer circle. Here the boundary data is obtained as $\mathbf{g} = \nabla u_{\text{ex}} \cdot \mathbf{n} - iku_{\text{ex}}$ from the exact solution

$$u_{\text{ex}}(x, y) = -\frac{J'_0(ka)}{H'_0(ka)} H_0^1(kr) - 2 \sum_{n=1}^{\infty} i^n \frac{J'_n(ka)}{H'_n(ka)} H_n^1(kr) \cos(n\theta), \quad (34)$$

where (r, θ) are polar coordinates of (x, y) , and J_n and H_n^1 are the Bessel function and the Hankel function of the first kind, respectively [31]. Note that the prime denotes the derivative of a function with respect to its argument.

This example serves to illustrate the ability of the phase-based HDG method to deal with diffractions at smooth convex boundaries. When an incident field hits a smooth obstacle there will be a *shadow region* behind it and the GO solution is discontinuous along the *shadow boundary* dividing the shadow part and the illuminated part of the obstacle. At the point (or curve in three dimensions) where the incident rays are tangent to the obstacle surface, the incident rays act as a source for surface rays, or *creeping rays*, that propagate along geodesics on the surface. The amplitude of the creeping rays is proportional to the amplitude of the incident rays and decays exponentially along the creeping ray's trajectory. At each point on a convex surface the creeping rays emit *surface-diffracted rays* in the tangential direction, which follow the usual

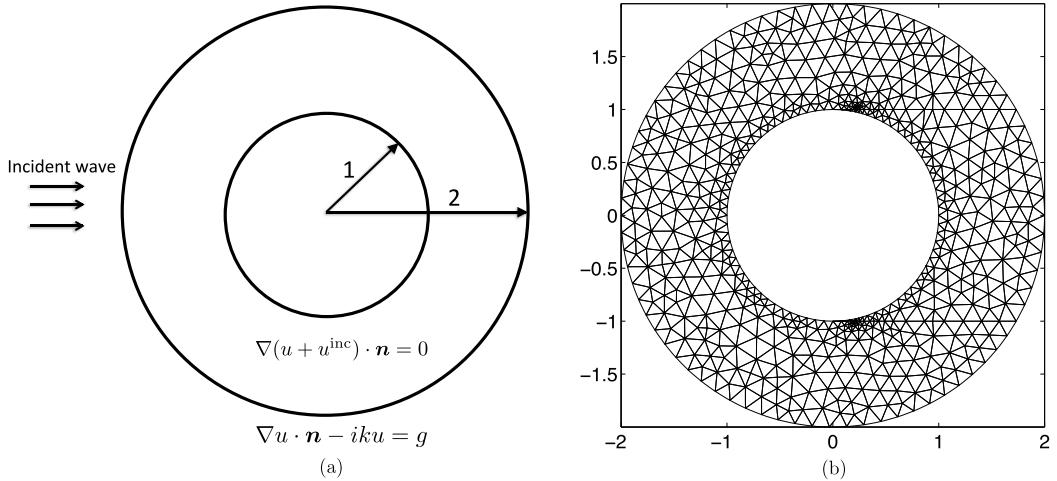


Fig. 2. Problem specification of the unit circle scatterer: (a) geometry and boundary conditions and (b) finite element mesh of 1136 triangular elements.

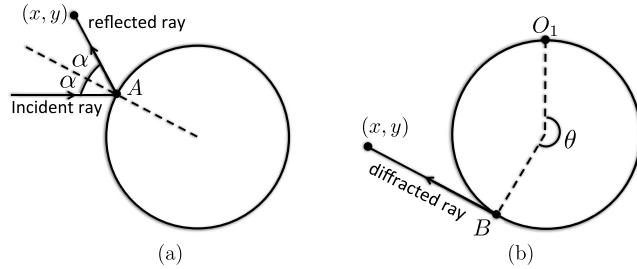


Fig. 3. An illustration of the ray tracing scheme for phase computation: (a) reflected ray and (b) diffracted ray.

GO laws. To compute the contribution of the creeping rays to the wave field, one thus needs to find these surface-diffracted waves.

To obtain the phases, we decompose the domain Ω such that $\bar{\Omega} = \bar{\Omega}^1 \cup \bar{\Omega}^2$, where $\Omega^1 = \{(x, y) \in \Omega : x > 0, |y| < 1\}$ (shadow region) and $\Omega^2 = \Omega \setminus \Omega^1$ (illuminated region). Both subdomains have three phases which are obtained analytically as follows. The first phase is formed by the incident rays in Ω^1 and the reflected rays in Ω^2 . In the subdomain Ω^1 , we know that $\varphi_1(x, y) = x$ since $u^{\text{inc}} = \exp(ikx)$. For any point $(x, y) \in \bar{\Omega}^2$, we use the law of reflection to compute $\varphi_1(x, y)$ as

$$\varphi_1(x, y) = x_A + \sqrt{(x - x_A)^2 + (y - y_A)^2}, \quad (35)$$

where, as shown in Fig. 3(a), $A = (x_A, y_A)$ is a point at which the incident ray hits the circle and generates a reflected ray that goes through (x, y) . The second phase $\varphi_2(x, y)$ corresponds to the creeping rays emanating from $O_1 = (0, 1)$. For any point $P = (x, y) \in \bar{\Omega}$ it is computed as

$$\varphi_2(x, y) = \theta + \sqrt{(x - x_B)^2 + (y - y_B)^2}, \quad (36)$$

where, as shown in Fig. 3(b), θ is the angle measured in radians between O_1 and B , and B is a point at which the vector \vec{BP} is tangent to the circle at B and directed in the clockwise direction. Note that the second phase φ_2 is discontinuous along the line $y = 1$ with $x \geq 0$ because θ is double-valued, namely $\theta = 0$ and $\theta = 2\pi$, at the point O_1 . The third phase $\varphi_3(x, y)$ corresponds to the creeping rays emanating from $O_2 = (0, -1)$ and is computed in a similar way as the second phase. Fig. 4 depicts the three phases φ_1 , φ_2 , and φ_3 .

Fig. 5 shows the plots of the approximate solution u_h for $k = 50, 100$ and 200 when we take $p = 2$ and $\epsilon = 10^{-5}$. We observe that the phase-based HDG method is capable of providing full-wave solutions at high frequencies on a very coarse mesh of 1136 triangular elements. We present in Fig. 6 the relative error measured in the $L^2(\Omega)$ norm and the condition number of the matrix system as a function of k . It is interesting to see that the relative error initially increases with k and then saturates around 6×10^{-3} for $k > 100$, while the condition number initially decreases with k and then fluctuates around 7×10^5 for $k > 120$. This is expected because the amplitudes are slowly varying functions in space and because the basis functions of the local spaces become rapidly oscillatory as k increases. These results clearly demonstrate that the proposed method allows us to achieve a given error tolerance at a cost which is almost independent of the wave number.

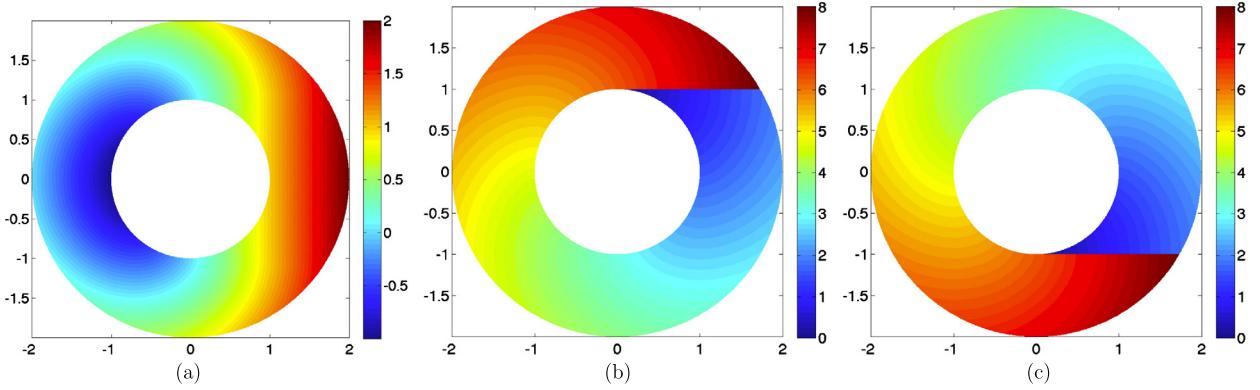


Fig. 4. Three phases for the scattering of an incident wave $u^{inc} = \exp(ikx)$ by the unit circle: (a) the first phase due to the reflected rays and the incident rays, (b) the second phase due to the diffracted rays emanating from $(0, 1)$, and (c) the third phase due to the diffracted rays emanating from $(0, -1)$. Note that the diffracted phases are discontinuous along the shadow boundary lines.

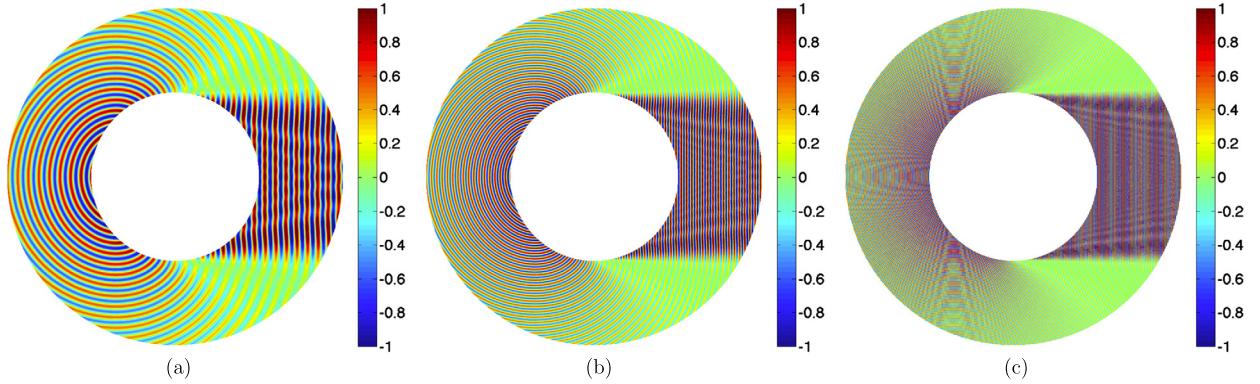


Fig. 5. Numerical results for the unit circle scatterer for $p = 2$: plots of u_h for (a) $k = 50$, (b) $k = 100$ and (c) $k = 200$.

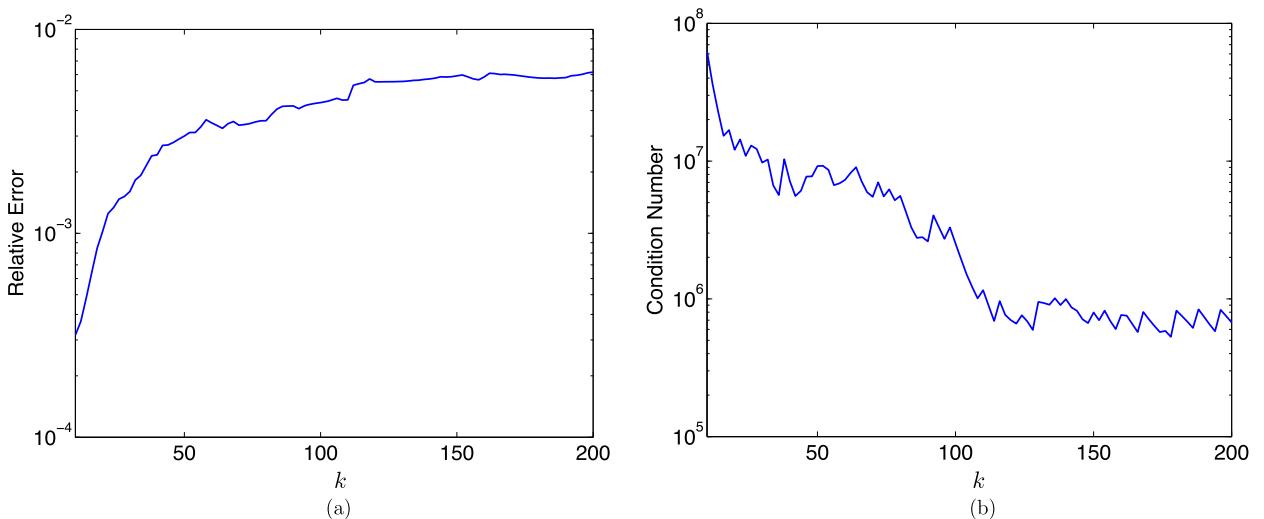


Fig. 6. Numerical results for the unit circle scatterer for $p = 2$: (a) relative error in the $L^2(\Omega)$ norm $\|u - u_h\|_{L^2(\Omega)} / \|u\|_{L^2(\Omega)}$ and (b) condition number of the global matrix as a function of k .

Table 2

Comparison between the phase-based HDG method and the CG-FEM, and the standard HDG method. Here N_e is the number of elements, N_{dof} is the number of global unknowns, and $E_{\text{rel}} = \|u - u_h\|_\Omega / \|u\|_\Omega$ denotes the relative error measured in the L^2 norm of the approximate solution.

k	Phase-based HDG			CG-FEM ($p = 2$)			Standard HDG ($p = 2$)		
	N_e	N_{dof}	E_{rel}	N_e	N_{dof}	E_{rel}	N_e	N_{dof}	E_{rel}
25	1136	8560	1.40E–03	27 270	55 140	2.55E–03	9780	44 550	3.08E–03
50	1136	9872	3.00E–03	109 208	219 616	4.38E–03	39 432	178 524	4.39E–03
100	1136	11 608	4.39E–03	436 176	874 752	8.24E–03	157 948	712 926	7.53E–03
150	1136	12 509	5.92E–03	983 356	1 970 312	1.21E–02	354 588	1 598 886	1.10E–02
200	1136	13 029	6.17E–03	1 751 518	3 507 836	1.61E–02	630 104	2 839 788	1.46E–02

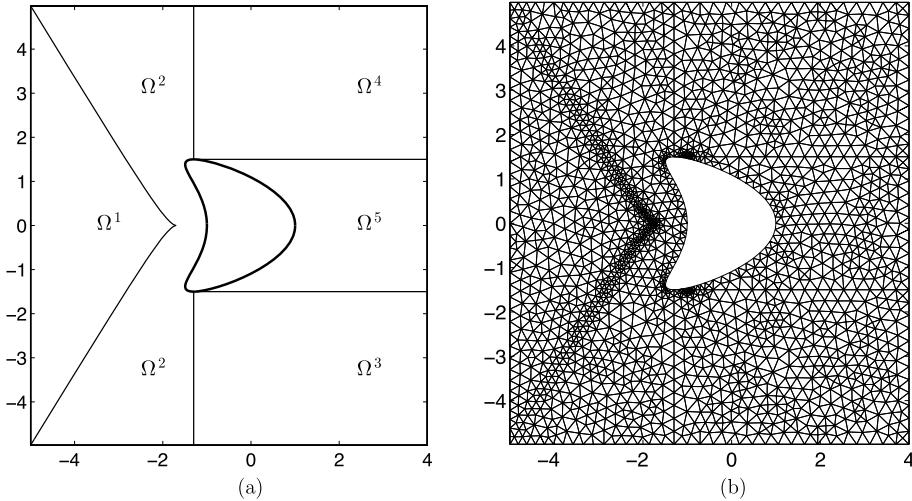


Fig. 7. Problem specification of the kite-shaped scatterer: (a) geometry and domain decomposition and (b) finite element mesh of 3983 triangular elements.

In Table 2 we compare the phase-based HDG method with the continuous Galerkin finite element method (CG-FEM) and the standard HDG method in terms of the global unknowns and relative errors for $k = 25, 50, 100, 150$, and 200. Here all the methods use polynomials of degree $p = 2$ to represent the approximate solution. Again we use a tolerance of $\epsilon = 10^{-5}$ in the orthogonalization process of the phase-based HDG method. We observe that the global unknowns of the CG-FEM and the standard HDG method increase very rapidly with k , while those of the phase-based HDG method grow very slowly with k . (Note that because the basis functions in (15) become more rapidly oscillatory as k increases, the orthogonalization process removes less redundant degrees of freedom as k increases. As a result, even though the mesh is the same, the number of global unknowns of the phase-based HDG method increases with k .) We also see that the phase-based HDG method yields smaller relative errors than the CG-FEM and the standard HDG method. Specifically, the phase-based HDG method provides a relative error of 6.17×10^{-3} with only 13 029 degrees of freedom for $k = 200$, while the CG-FEM (the standard HDG method) produces a relative error of 1.61×10^{-2} (1.46×10^{-2}) with 3 507 836 (2 839 788) degrees of freedom. Hence, the phase-based HDG method results in a degree-of-freedom saving of two orders of magnitude relative to both the CG-FEM and the standard HDG method. Clearly, this saving is made possible only through the introduction of the GO phases into the HDG method.

Finally, we emphasize that the orthogonalization is crucial to ensure not only the efficiency but also the stability of the phase-based HDG method. Indeed, since the first phase and the second phase have the same value along the shadow boundary line $y = 1, x \geq 0$ (and the first phase and the third phase have the same value along the shadow boundary line $y = -1, x \geq 0$), the phase-based HDG method without orthogonalization becomes ill-conditioned and unstable.

4.3. Sound-hard scattering from a kite-shaped scatterer

We consider the computation of a wave field scattering from a kite-shaped scatterer under an incident plane wave $u^{\text{inc}} = \exp(ikx)$. The kite-shaped scatterer is represented by the following parameterized curve

$$x = \cos(t) + 0.65 \cos(2t) - 0.65, \quad y = 1.5 \sin(t), \quad t \in [0, 2\pi]. \quad (37)$$

Fig. 7 shows the computational domain and finite element mesh. The Neumann condition $\nabla(u + u^{\text{inc}}) \cdot \mathbf{n} = 0$ is imposed on the kite boundary, while the Robin condition $\nabla u \cdot \mathbf{n} - iku = 0$ on the outer boundary which is the boundary of a rectangular domain $(-5, 4) \times (-4.9851, 4.9851)$.

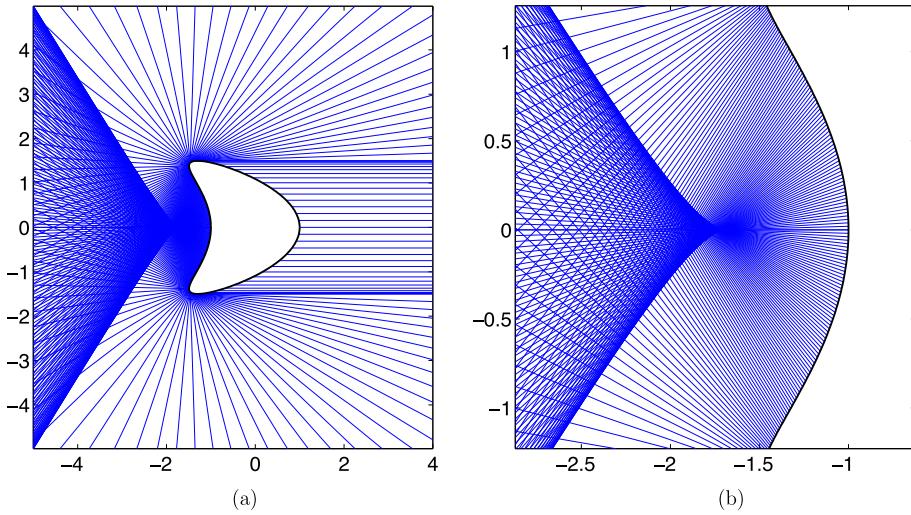


Fig. 8. The fold caustic generated by the reflection of an incident plane wave $u^{\text{inc}}(x, y) = \exp(ikx)$ from the kite-shaped scatterer and computed by ray tracing: (a) the reflected rays in the illuminated region and the incident rays in the shadow region, (b) and the zoom-up view near the cusp singularity at $(-1.7020, 0)$.

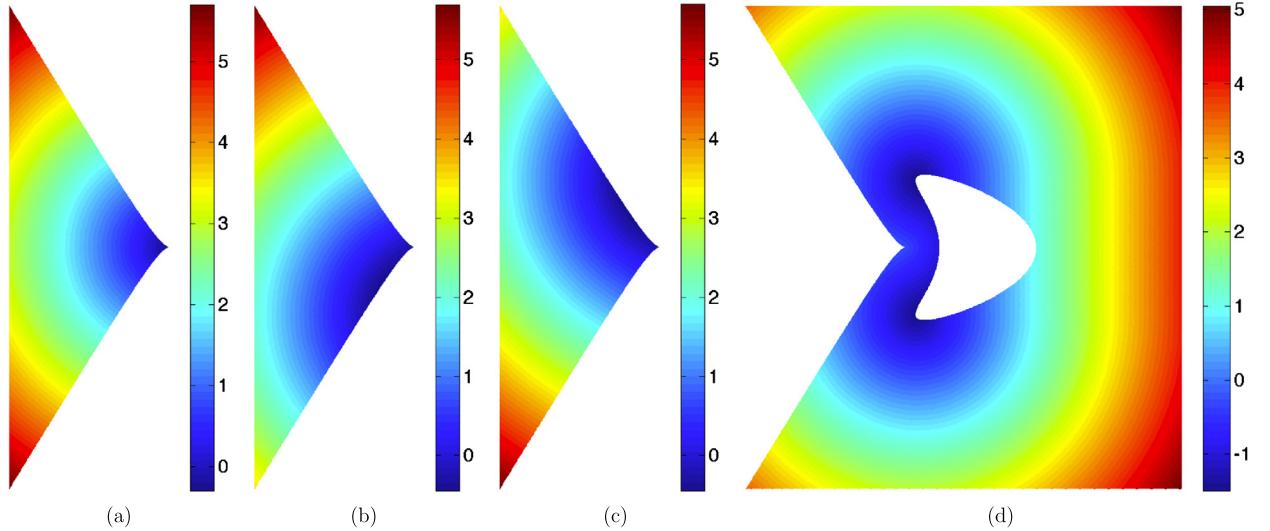


Fig. 9. The phases correspond to the ray families shown in Fig. 8(a): (a) the first phase φ_1^1 in Ω^1 , (b) the second phase φ_2^1 in Ω^1 , (c) the third phase φ_3^1 in Ω^1 , and (d) the phases $\varphi_1^s, s = 2, 3, 4, 5$ in the other subdomains.

This example serves to illustrate the ability of the proposed method to deal with *caustics* due to reflection from smooth concave boundaries. Fig. 8(a) depicts the rays reflected by the kite-shaped scatterer in the illuminated region and the incident rays in the shadow region. Due to the concavity of the scatterer, multiple families of rays reflect from the kite boundary to form a *fold caustic* with a *cusp singularity* at the point $(-1.7020, 0)$ as shown in Fig. 8(b). The domain is divided into 5 subdomains as shown in Fig. 7: Ω^1 is the caustic region, Ω^2 is the illuminated region close to the caustic region, Ω^3 and Ω^4 are the illuminated regions close to the shadow region, and Ω^5 is the shadow region. Unlike the previous examples in which the phases are determined analytically, we compute the phases by using ray tracing and solving the eikonal equations. In particular, we obtain the three phases in Fig. 9(a), (b), and (c) by numerically solving

$$|\nabla \varphi_n^1| = 1 \quad \text{in } \Omega^1, \quad \varphi_n^1 = r_n^1 \quad \text{on } \partial\Omega^1, \quad n = 1, 2, 3, \quad (38)$$

where r_n^1 is the value of the n th phase on $\partial\Omega^1$ and computed by means of ray tracing as shown in Fig. 8(a). Likewise, we compute the phase in Fig. 9(d) by numerically solving

$$|\nabla \varphi_1^s| = 1 \quad \text{in } \Omega^s, \quad \varphi_1^s = r_1^s \quad \text{on } \partial\Omega^s, \quad s = 2, 3, 4, 5, \quad (39)$$

where the boundary values r_1^s are also computed by means of ray tracing.

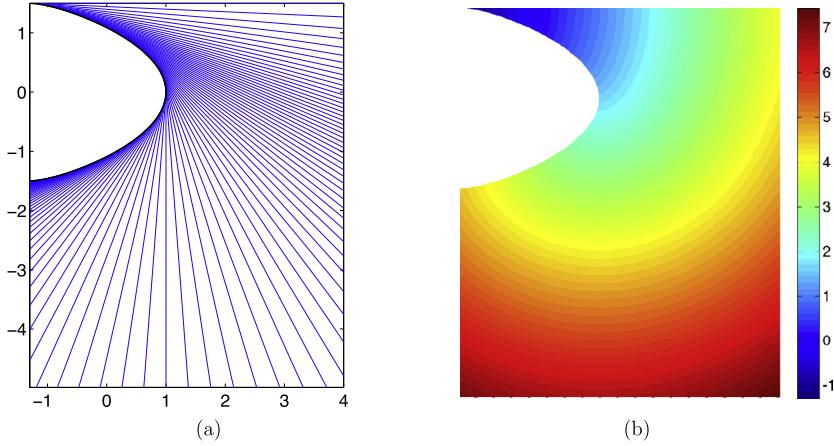


Fig. 10. The creeping rays (a) and the associated phases (b) emanate from $(-1.3, 1.5)$.

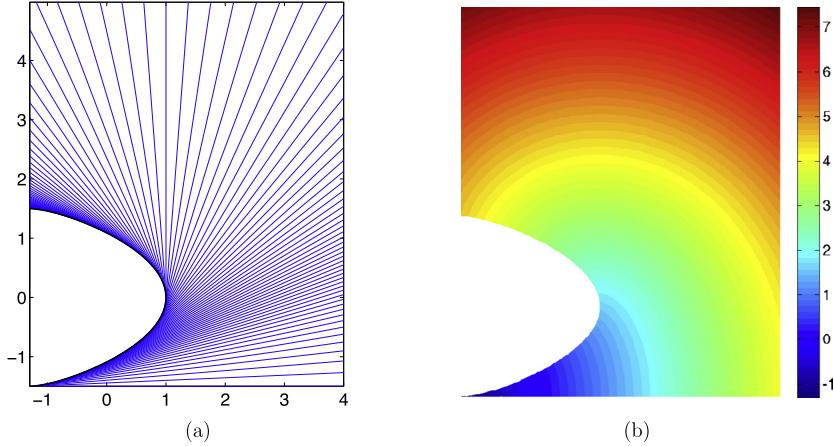


Fig. 11. The creeping rays (a) and the associated phases (b) emanate from $(-1.3, -1.5)$.

Next, we compute the phases due to diffractions at smooth boundaries. The phases in Fig. 10(b) are obtained by numerically solving

$$\begin{aligned} |\nabla \varphi_2^5| &= 1 \quad \text{in } \Omega^5, & \varphi_2^5 = r_2^5 & \quad \text{on } \partial\Omega^5, \\ |\nabla \varphi_2^3| &= 1 \quad \text{in } \Omega^3, & \varphi_2^3 = r_2^3 & \quad \text{on } \partial\Omega^3, \end{aligned} \quad (40)$$

where, shown in Fig. 10(a), r_2^5 and r_2^3 are computed by means of ray tracing. The phases in Fig. 11(b) are obtained by numerically solving

$$\begin{aligned} |\nabla \varphi_3^5| &= 1 \quad \text{in } \Omega^5, & \varphi_3^5 = r_3^5 & \quad \text{on } \partial\Omega^5, \\ |\nabla \varphi_2^4| &= 1 \quad \text{in } \Omega^4, & \varphi_2^4 = r_2^4 & \quad \text{on } \partial\Omega^4, \end{aligned} \quad (41)$$

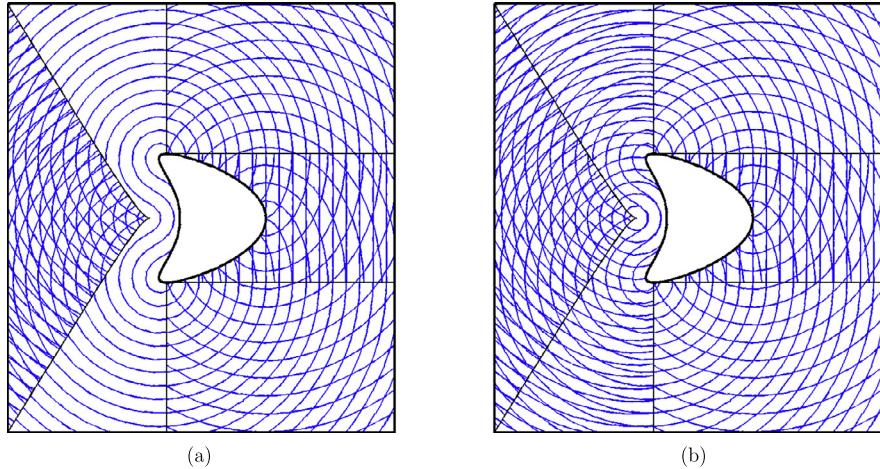
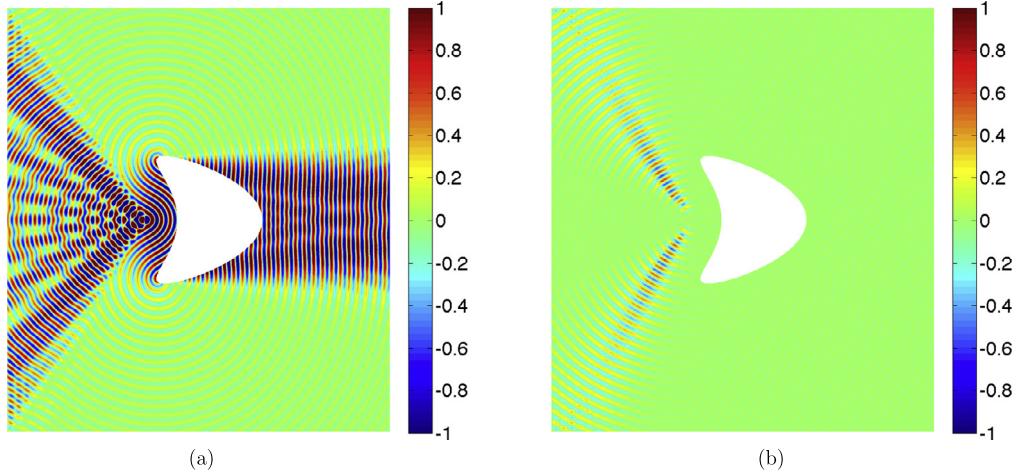
where, shown in Fig. 11(a), r_3^5 and r_2^4 are computed by means of ray tracing.

We present in Fig. 12(a) the contour plot of the phases computed above and summarize in Table 3 (the middle column) the number of phases on 5 subdomains. Note that while Ω^1 consists of 3 phases, Ω^2 has only one phase and that there are two different phases on the caustic curve $\Gamma_C \equiv \overline{\Omega^1} \cap \overline{\Omega^2}$. This causes a poor approximation of the wave field along the caustic curve. To see this issue, we present in Fig. 13(a) the scattered field for $k = 30$, which is computed by the phase-based HDG method with $p = 2$ and $\epsilon = 10^{-5}$. We further depict in Fig. 13(b) the difference between this solution and a reference solution computed by the standard HDG method on the same mesh with polynomial degree $p = 16$. We see that the difference is quite significant near the caustic curve. This result indicates the solution obtained with the phase-based HDG method is inaccurate near the caustic curve. This is because the phase system in Fig. 12(a) is not adequate to resolve the wave field in the subdomain Ω^2 . As demonstrated next, this issue can be remedied by introducing an additional phase into the subdomain Ω^2 .

Table 3

Number of phases on 5 subdomains for the two collections of phases presented in Fig. 12.

Subdomain	Original phases in Fig. 12(a)	Enhanced phases in Fig. 12(b)
Ω^1	3 phases	3 phases
Ω^2	1 phases	2 phases
Ω^3	2 phases	2 phases
Ω^4	2 phases	2 phases
Ω^5	3 phases	3 phases

**Fig. 12.** Contour plot of the phases for (a) the original phase structure and (b) the enhanced phase structure.**Fig. 13.** Numerical solution of the kite-shaped scatterer for $k = 30$: the scattered field computed using the phase structure in Fig. 12(a) and the difference between this scattered field and a reference solution. Note that the difference is quite significant in the caustic transition region.

In particular, we extend the phase shown in Fig. 9(a) from subdomain Ω^1 into subdomain Ω^2 . We note that the extension of the phase in Fig. 9(a) into Ω^2 can not be done by means of ray tracing because its rays do not follow the law of reflection. However, this phase must still satisfy the eikonal equation in Ω^2 and its value is known on the caustic curve. Therefore, we compute the phase by solving the following eikonal problem

$$|\nabla \varphi_2^2| = 1 \quad \text{in } \Omega^2, \quad \varphi_2^2 = r_2^2 \quad \text{on } \Gamma_C \equiv \bar{\Omega}^1 \cap \bar{\Omega}^2, \quad (42)$$

where r_2^2 is the value of the phase in Fig. 9(a) on the caustic curve. The resulting phase structure is shown in Fig. 12(b) and summarized in the last column of Table 3. Using this new phase structure, we compute the scattered field (for $p = 2$ and $\epsilon = 10^{-5}$) and plot the results in Fig. 14. We see that the difference between the resulting solution and the reference solution is small in the entire domain. Hence, the additional phase φ_2^2 is crucial for the treatment of the fold caustic. Without it, we would obtain much less accurate solutions as illustrated in Fig. 13.

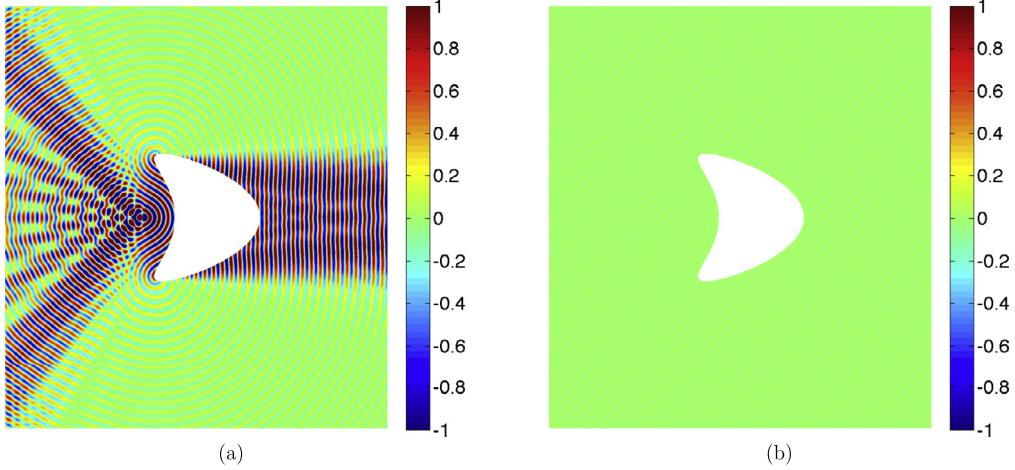


Fig. 14. Numerical solution of the kite-shaped scattering problem for $k = 30$: the scattered field computed using the phase collection in Fig. 12(b) and the difference between this scattered field and a reference solution. Note that the difference is now small in the entire domain.

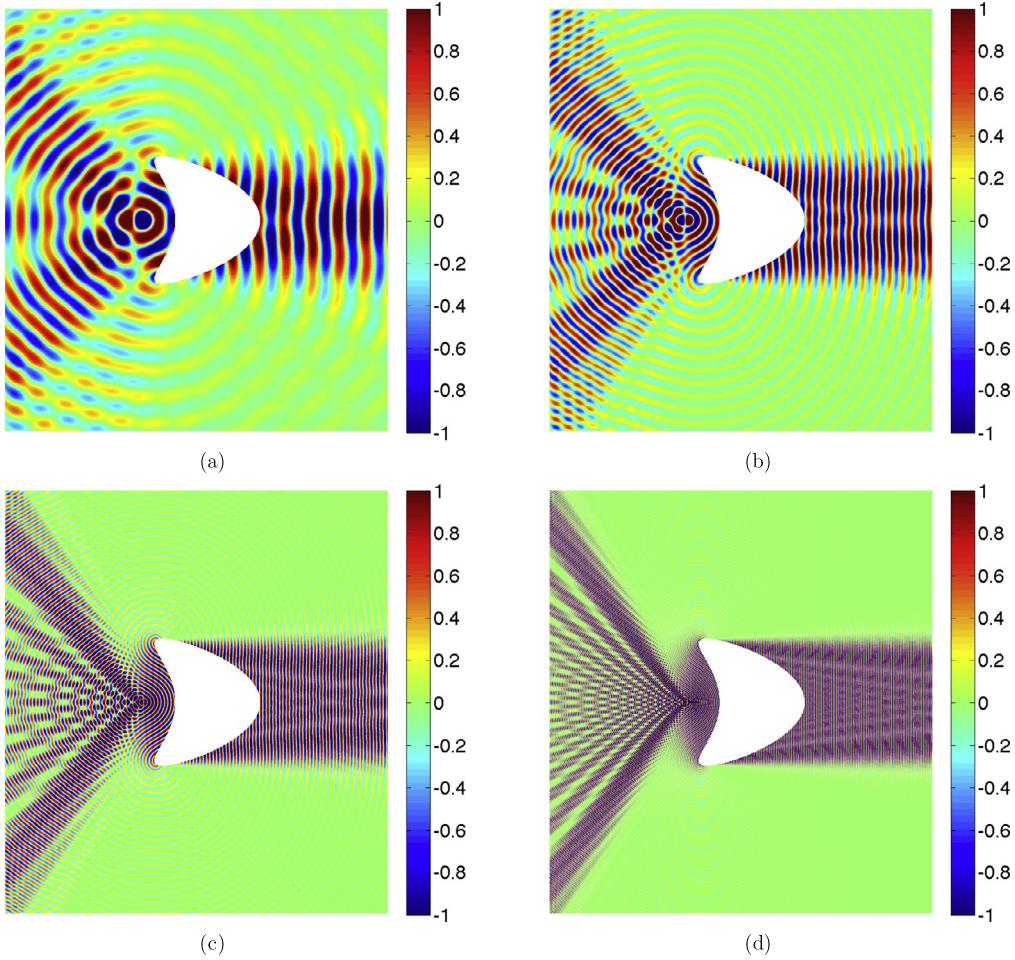


Fig. 15. The scattered field computed using the phase collection in Fig. 12(b) for: (a) $k = 10$, (b) $k = 20$, (c) $k = 40$, and (d) $k = 80$.

Finally, we present in Fig. 15 the scattered field computed using the collection of phases in Fig. 12(b) for various values of k . These results are computed on the same mesh in Fig. 12(b) with $p = 2$ and $\epsilon = 10^{-5}$. We see that the phase-based HDG method yields well-resolved solutions on the same mesh even though the wavenumber increases rapidly. On the other

hand, the standard HDG method requires a finite element mesh of 939 850 elements with $p = 2$ to obtain a well-resolved solution for $k = 80$. Note that the CG-FEM needs a considerably finer mesh than the standard HDG method in order to achieve a similar resolution.

5. Conclusions

In this paper, we introduce a phase-based HDG method for the numerical solution of the Helmholtz equation. The method combines the powerful ideas of geometrical optics and HDG methodologies. More precisely, the computational procedure of the proposed method is decomposed into two stages. In the first stage, we develop a new scheme that combines the HDG method with ray tracing to compute multivalued solution of the eikonal equation. In the second stage, we first embed these phases with standard polynomials to construct the local approximation spaces that are capable of approximating highly oscillatory solutions of the Helmholtz equation; we then orthogonalize the original basis functions to obtain orthogonal basis functions, thereby increasing the numerical stability and reducing redundant degrees of freedom; and finally, we implement the phase-based HDG method with an appropriate choice of the stabilization parameter. We present numerical results to demonstrate the efficiency of the method and compare its performance with that of the CG-FEM and the standard HDG method. Numerical results show that the method is capable of providing well-resolve solutions for a wide range of frequencies on the same coarse mesh. Moreover, they also show that the number of unknowns required to achieve a desire accuracy with the phase-based HDG method is several orders of magnitude less than those with the CG-FEM and the standard HDG method. The phase-based HDG method is ideally suited to applications that require fast frequency sweep since the first stage is performed only once.

We conclude the paper by pointing out possible extensions and directions for future research. We would like to extend the present method to the time-harmonic Maxwell's equations as well as the elastic wave equations. The *a priori* and *a posteriori* error analysis of the present method constitute subjects on ongoing research, as the method will greatly benefit from h/p -mesh adaptivity as corners, edges, caustics, and shadow boundaries which obviously require more mesh resolutions than other regions of the domain.

Acknowledgements

N.C. Nguyen and J. Peraire would like to acknowledge the partial support by AFOSR Grant No. FA9550-11-1-0141, AFOSR Grant No. FA9550-12-0357, and the Singapore-MIT Alliance for Research and Technology Centre. B. Cockburn was partially supported by the National Science Foundation (Grant DMS-0712955) and by the Minnesota Supercomputing Institute.

References

- [1] D. Andersh, J. Moore, S. Kosanovich, D. Kapp, R. Bhalla, R. Kipp, T. Courtney, A. Nolan, F. German, J. Cook, J. Hughes, Xpatch 4: the next generation in high frequency electromagnetic modeling and simulation software, in: Record of the IEEE 2000 International Radar Conference [Cat. No. 00CH37037], IEEE, 2000, pp. 844–849.
- [2] I. Babuška, J.M. Melenk, The partition of unity method, *Int. J. Numer. Methods Eng.* 40 (1997) 727–758.
- [3] P.E. Barbone, J.M. Montgomery, O. Michael, I. Harari, Scattering by a hybrid asymptotic/finite element method, *Comput. Methods Appl. Mech. Eng.* 164 (1998) 141–156.
- [4] J.D. Benamou, Big ray tracing: multivalued travel time field computation using viscosity solutions of the eikonal equation, *J. Comput. Phys.* 474 (1996) 463–474.
- [5] D. Bouche, F. Molinet, R. Mittra, *Asymptotic Methods in Electromagnetics*, Springer, 1997.
- [6] O.P. Bruno, C.A. Geuzaine, J.A. Monro, F. Reitich, Prescribed error tolerances within fixed computational times for scattering problems of arbitrarily high frequency: the convex case, *Philos. Trans. R. Soc., Math. Phys. Eng. Sci.* 362 (1816) (March 2004) 629–645.
- [7] F. Celiker, B. Cockburn, K. Shi, Hybridizable discontinuous Galerkin methods for Timoshenko beams, *J. Sci. Comput.* 44 (1) (March 2010) 1–37.
- [8] 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) (February 1998) 255–299.
- [9] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics in Applied Mathematics, vol. 40, SIAM, Amsterdam, 2002.
- [10] B. Chabaud, B. Cockburn, Uniform-in-time superconvergence of HDG methods for the heat equation, *Math. Comput.* 81 (2012) 107–129.
- [11] B. Cockburn, B. Dong, J. Guzmán, A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems, *Math. Comput.* 77 (2008) 1887–1916.
- [12] B. Cockburn, B. Dong, J. Guzmán, M. Restelli, R. Sacco, A hybridizable discontinuous Galerkin method for steady-state convection-diffusion-reaction problems, *SIAM J. Sci. Comput.* 31 (5) (2009) 3827–3846.
- [13] B. Cockburn, J. Gopalakrishnan, R. Lazarov, Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems, *SIAM J. Numer. Anal.* 47 (2) (January 2009) 1319–1365.
- [14] B. Cockburn, J. Gopalakrishnan, N.C. Nguyen, J. Peraire, F.-J. Sayas, Analysis of HDG methods for Stokes flow, *Math. Comput.* 80 (2011) 723–760.
- [15] B. Cockburn, J. Gopalakrishnan, F.-J. Sayas, A projection-based error analysis of HDG methods, *Math. Comput.* 79 (2010) 1351–1367.
- [16] B. Cockburn, J. Guzmán, H. Wang, Superconvergent discontinuous Galerkin methods for second-order elliptic problems, *Math. Comput.* 78 (2009) 1–24.
- [17] B. Cockburn, N.C. Nguyen, J. Peraire, A comparison of HDG methods for Stokes flow, *J. Sci. Comput.* 45 (1–3) (March 2010) 215–237.
- [18] B. Cockburn, J. Qian, F. Reitich, J. Wang, An accurate spectral/discontinuous finite-element formulation of a phase-space-based level set approach to geometrical optics, *J. Comput. Phys.* 208 (1) (September 2005) 175–195.
- [19] B. Cockburn, V. Quenneville-Bélair, Uniform-in-time superconvergence of HDG methods for the acoustic wave equation, *Math. Comput.* 83 (2014) 65–85.
- [20] B. Cockburn, F.J. Sayas, Divergence-conforming HDG methods for Stokes flow, *Math. Comput.* 83 (2014) 1571–1598.
- [21] B. Cockburn, C.-W. Shu, Runge–Kutta discontinuous Galerkin methods for convection-dominated problems, *J. Sci. Comput.* 16 (2001) 173–261.
- [22] J. Cui, W. Zhang, An analysis of the HDG method for the Helmholtz equation, *IMA J. Numer. Anal.* 34 (2014) 279–295.

- [23] E. Ekelman, G. Thiele, A hybrid technique for combining the moment method treatment of wire antennas with the GTD for curved surfaces, *IEEE Trans. Antennas Propag.* 28 (6) (November 1980) 831–839.
- [24] B. Engquist, O. Runborg, Computational high frequency wave propagation, *Acta Numer.* 12 (May 2003) 181–266.
- [25] C. Farhat, I. Harari, U. Hetmaniuk, A discontinuous Galerkin method with Lagrange multipliers for the solution of Helmholtz problems in the mid-frequency regime, *Comput. Methods Appl. Mech. Eng.* 192 (2003) 1389–1419.
- [26] E. Fatemi, B. Engquist, S. Osher, Numerical solution of the high frequency asymptotic expansion for the scalar wave equation, *J. Comput. Phys.* 120 (1995) 145–155.
- [27] X. Feng, Y. Xing, Absolutely stable local discontinuous Galerkin methods for the Helmholtz equation with large wave number, *Math. Comput.* 82 (283) (October 2012) 1269–1296.
- [28] G. Gabard, Discontinuous Galerkin methods with plane waves for time-harmonic problems, *J. Comput. Phys.* 225 (2) (August 2007) 1961–1984.
- [29] E. Giladi, J.B. Keller, A hybrid numerical asymptotic method for scattering problems, *J. Comput. Phys.* 174 (1) (November 2001) 226–247.
- [30] R. Griesmaier, P. Monk, Error analysis for a hybridizable discontinuous Galerkin method for the Helmholtz equation, *J. Sci. Comput.* 49 (3) (2011) 291–310.
- [31] Isaa Harari, Thomas J.R. Hughes, Galerkin/least-squares finite element methods for the reduced wave equation with non-reflecting boundary conditions in unbounded domains, *Comput. Methods Appl. Mech. Eng.* 98 (3) (August 1992) 411–454.
- [32] R. Hiptmair, A. Moiola, I. Perugia, Plane wave discontinuous Galerkin methods for the 2D Helmholtz equation: analysis of the p-version, *SIAM J. Numer. Anal.* 49 (1) (2011) 264–284.
- [33] L.N.T. Huynh, N.C. Nguyen, J. Peraire, B.C. Khoo, A high-order hybridizable discontinuous Galerkin method for elliptic interface problems, *Int. J. Numer. Methods Eng.* 93 (2) (January 2013) 183–200.
- [34] S. Langdon, S.N. Chandler-Wilde, A wavenumber independent boundary element method for an acoustic scattering problem, *SIAM J. Numer. Anal.* 43 (6) (January 2006) 2450–2477.
- [35] D.A. McNamara, C.W.I. Pistoia, *Introduction to the Uniform Geometrical Theory of Diffraction* (Artech House Microwave Library), Artech Print on Demand, 1990.
- [36] D. Moro, N.C. Nguyen, J. Peraire, Navier–Stokes solution using hybridizable discontinuous Galerkin methods, in: 20th AIAA Computational Fluid Dynamics Conference, American Institute of Aeronautics and Astronautics, Honolulu, Hawaii, June 2011, AIAA-2011-3407.
- [37] N.C. Nguyen, J. Peraire, Hybridizable discontinuous Galerkin methods for partial differential equations in continuum mechanics, *J. Comput. Phys.* 231 (18) (July 2012) 5955–5988.
- [38] N.C. Nguyen, J. Peraire, B. Cockburn, An implicit high-order hybridizable discontinuous Galerkin method for linear convection diffusion equations, *J. Comput. Phys.* 228 (9) (May 2009) 3232–3254.
- [39] N.C. Nguyen, J. Peraire, B. Cockburn, An implicit high-order hybridizable discontinuous Galerkin method for nonlinear convection diffusion equations, *J. Comput. Phys.* 228 (23) (December 2009) 8841–8855.
- [40] N.C. Nguyen, J. Peraire, B. Cockburn, A hybridizable discontinuous Galerkin method for Stokes flow, *Comput. Methods Appl. Mech. Eng.* 199 (9–12) (January 2010) 582–597.
- [41] N.C. Nguyen, J. Peraire, B. Cockburn, An implicit high-order hybridizable discontinuous Galerkin method for the incompressible Navier–Stokes equations, *J. Comput. Phys.* 230 (4) (February 2011) 1147–1170.
- [42] N.C. Nguyen, J. Peraire, B. Cockburn, High-order implicit hybridizable discontinuous Galerkin methods for acoustics and elastodynamics, *J. Comput. Phys.* 230 (10) (May 2011) 3695–3718.
- [43] N.C. Nguyen, J. Peraire, B. Cockburn, Hybridizable discontinuous Galerkin methods for the time-harmonic Maxwell's equations, *J. Comput. Phys.* 230 (19) (August 2011) 7151–7175.
- [44] S. Osher, L.-T. Cheng, M. Kang, H. Shim, Y.-H. Tsai, Geometric optics in a phase-space-based level set and Eulerian framework, *J. Comput. Phys.* 179 (2) (July 2002) 622–648.
- [45] J. Peraire, N.C. Nguyen, B. Cockburn, A hybridizable discontinuous Galerkin method for the compressible Euler and Navier–Stokes equations, in: 48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, June, 2010, number AIAA-2010-363.
- [46] B. Pluymers, B. Hal, D. Vandepitte, W. Desmet, Trefftz-based methods for time-harmonic acoustics, *Arch. Comput. Methods Eng.* 14 (4) (August 2007) 343–381.
- [47] F. Reitich, K.K. Tamma, State-of-the-art, trends, and directions in computational electromagnetics, *Comput. Model. Eng. Sci.* 5 (4) (2004) 287–294.
- [48] O. Runborg, Mathematical models and numerical methods for high frequency waves, *Commun. Comput. Phys.* 2 (5) (2007) 827–880.
- [49] J.A. Sethian, *Level Set Methods and Fast Marching Methods*, Cambridge University Press, Cambridge, 1999.
- [50] S.-C. Soon, B. Cockburn, H.K. Stolarski, A hybridizable discontinuous Galerkin method for linear elasticity, *Int. J. Numer. Methods Eng.* 80 (8) (2009) 1058–1092.
- [51] W.W. Symes, J. Qian, A slowness matching Eulerian method for multivalued solutions of eikonal equations, *J. Sci. Comput.* 19 (December 2003) 501–526.
- [52] J. Van Trier, W.W. Symes, Upwind finite-difference calculation of traveltimes, *Geophysics* 56 (1991) 812–821.
- [53] J. Vidale, Finite-difference calculation of travel times, *Bull. Seismol. Soc. Am.* 78 (6) (December 1988) 2062–2076.