A comparison of high-order polynomial and wave-based methods for Helmholtz problems

Alice Lieu\textsuperscript{a}, Gwénaël Gabard\textsuperscript{a}, Hadrien Bériot\textsuperscript{b}

\textsuperscript{a}Institute of Sound and Vibration Research, University of Southampton, Southampton, SO17 1BJ, United Kingdom
\textsuperscript{b}Siemens PLM, Interleuvenlaan 68, Researchpark Z1, Leuven, Belgium

Abstract

The application of computational modelling to wave propagation problems is hindered by the dispersion error introduced by the discretisation. Two common strategies to address this issue are to use high-order polynomial shape functions (e.g. $hp$-FEM), or to use physics-based, or Trefftz, methods where the shape functions are local solutions of the problem (typically plane waves). Both strategies have been actively developed over the past decades and both have demonstrated their benefits compared to conventional finite-element methods, but they have yet to be compared. In this paper a high-order polynomial method ($p$-FEM with Lobatto polynomials) and the wave-based discontinuous Galerkin method are compared for two-dimensional Helmholtz problems. A number of different benchmark problems are used to perform a detailed and systematic assessment of the relative merits of these two methods in terms of interpolation properties, performance and conditioning. It is generally assumed that a wave-based method naturally provides better accuracy compared to polynomial methods since the plane waves or Bessel functions used in these methods are exact solutions of the Helmholtz equation. Results indicate that this expectation does not necessarily translate into a clear benefit, and that the differences in performance, accuracy and conditioning are more nuanced than generally assumed. The high-order polynomial method can in fact deliver comparable, and in some cases superior, performance compared to the wave-based DGM. In addition to benchmarking the intrinsic computational performance of these methods, a number of practical issues associated with realistic applications are also discussed.

Keywords: Comparison, High-order Finite Element Method, Wave-based Discontinuous Galerkin Method, Helmholtz problems

1. Introduction

Numerical dispersion is the main source of inaccuracy in the conventional finite element method when solving Helmholtz problems at high frequencies. For instance with linear elements the numerical error scales like $k^2 h$ with wavenumber $k$ and elements of size $h$. As $k$ increases, the number of elements per wavelength must also increase to maintain the same level of accuracy, which rapidly becomes prohibitively expensive. A number of different strategies have been explored to address this issue \cite{1,2}. The present paper provides a detailed comparison of two of the most common techniques: high-order polynomial methods and wave-based methods.

High-order finite element methods replace the standard, low-order Lagrange polynomials with higher-order functions providing superior interpolation properties. Different families of polynomials have been considered such as Bernstein, Hermite or Lobatto polynomials (comparisons of these can be found in \cite{3,4}). Fourier series or Tchebychev polynomials are also used with spectral methods \cite{5}. In this work, the framework of $p$-FEM is used to construct a continuous, high-order approximation with Lobatto shape functions for solutions of the Helmholtz equation \cite{6,7}. With this method the dispersion error is drastically reduced \cite{8} and it has been shown to be a valid approach to address the pollution effect and tackle large-scale problems \cite{9,10}. The $p$-FEM approach provides exponential convergence
When increasing the polynomial order $p$. It is also well-suited for $p$-adaptivity where the polynomial order is adjusted locally in the computational domain [11]. Finally, the hierarchic nature of the Lobatto shape functions leads to efficient algorithms for solving the same model over a range of frequencies [10].

Wave-based methods represent another category of methods developed to remedy the pollution problem (a detailed review of these methods can be found in [12]). The rationale is to incorporate a priori knowledge about the local behaviour of the solution into the numerical method. This is generally achieved by using local canonical solutions of the governing equations to build an approximation basis and these methods are therefore also called Trefftz methods. For the Helmholtz equation, the local solutions are generally plane waves, but other alternatives can be used, such as Bessel functions [13]. The central argument for using a basis of plane waves is that it follows directly from the governing equations and thus it allows embedding key features of the underlying physics into the numerical method. It is generally assumed that a plane-wave basis naturally provides better accuracy compared to polynomials since the latter bears no relation with the governing equations. One of the objectives of this paper is to assess this assumption.

A number of wave-based methods rely on discontinuous formulations where the solution is approximated with plane waves in each element and the continuity between elements is imposed weakly. This includes the Ultra Weak Variational Formulation (UWVF) [14] [15] [16], the least square method [17], and the wave-based discontinuous Galerkin method [18]. They differ in the way the inter-element continuity is formulated but it is now recognised that these methods are all variants of the same discontinuous Galerkin method using different numerical fluxes to ensure continuity between elements [19]. The DGM framework is particularly well suited for adaptivity since it is straightforward to change the number of plane waves in each element. A number of theoretical analyses have also been reported for this particular method [20] [21] [22] [23]. Results from [24] indicate that the wave-based DGM exhibits exponential convergence with respect to the number of degrees of freedom for very general element shapes, including in the presence of strong mesh refinements. This paves the way to for the development of fully automatic $hp$-adaptive versions of the method and a posteriori error estimators are being investigated [25]. In this work, the wave-based DGM from [18] is used. Numerical results in [19] [26] indicate that it is more efficient than the UWVF and least-square methods.

The Discontinuous Enrichment Method (DEM) is another example of methods using discontinuous interpolations [27]. Enrichment functions are added to the set of conventional shape functions to construct the approximate solution in each element. Lagrange multipliers defined on the edges between elements enforce the continuity of the solution. This method has been used to solve the Helmholtz equation [28] and extended to other fields such as fluid-structure interactions [29]. Removing the conventional shape functions from the DEM leads to a wave-based discontinuous Galerkin method with Lagrange multipliers [28]. The DEM leads to a significant reduction of the number of DOFs when compared to linear, quadratic and quartic Lagrange finite elements for a fixed accuracy in 2D [28] [30] and 3D [31].

Other Trefftz methods relying on discontinuous interpolations include the wave-based method [32] and the variational theory of complex rays (VTCR) [33], but these two methods cannot be formulated in the DG framework [12]. Results in [34] indicate that the UWVF tends to outperform the VTCR for simple Helmholtz problems.

The Partition of Unity FEM (PUFEM) constructs an approximation of the solution by multiplying enrichments with the standard FEM shape functions [35]. The benefit of this approach is that the resulting approximation space is naturally continuous. The downside is the calculation of the element matrices, which can be costly because it involves highly-oscillatory integrals with products of exponentials and polynomials. Elements of comparisons of PUFEM with the UWVF can be found in [36] [37] and the UWVF is relatively better conditioned and generally more efficient.

Both high-order polynomial FEM and wave-based methods have seen sustained developments over the past decades and have been applied to a wide range of disciplines (acoustics, aero-acoustics, linear elasticity, water waves, electromagnetism, porous materials). While partial elements of comparison exist in the literature [28] [30] [31], there hasn’t been a systematic and detailed assessment of their relative merits for Helmholtz problems. This is the focus of the present paper. Using a number of two-dimensional benchmark problems it provides a detailed comparison of the accuracy, cost and conditioning of two of these methods: $p$-FEM and the wave-based discontinuous Galerkin method.

The structure of the paper is as follows. Section 2 describes the two numerical methods and the benchmark problems are presented in Section 3. Section 4 discusses the different metrics used to assess the cost and accuracy of the numerical methods. The interpolation properties of each basis (polynomial and wave-based) are compared in Section 5. This is followed in Section 6 by the convergence and relative performance of the two methods to achieve a
given accuracy. Practical aspects of the two computational methods are then discussed in Section 7.

2. Presentation of the methods

We consider the propagation of harmonic, linear sound waves in a uniform, quiescent medium within a bounded domain Ω. Throughout this paper an implicit time dependence $e^{i\omega t}$ is assumed, where $\omega$ is the angular frequency. In the absence of volume sources, the complex amplitude $u$ of the acoustic pressure is governed by the homogeneous Helmholtz equation:

$$-k^2 u - \nabla^2 u = 0,$$

where $k = \omega/c_0$ is the free-field wavenumber and $c_0$ is the speed of sound. The boundary $\partial \Omega$ of the domain is equipped with an outward unit normal $n$. To generate a given solution $u_{ex}$ inside the computational domain an inhomogeneous Robin condition is applied on $\Gamma_r$:

$$\nabla u \cdot n + iku = g,$$

where the source term is given by $g = iku_{ex} + \nabla u_{ex} \cdot n$. On the remainder of the boundary $\Gamma_d = \partial \Omega \setminus \Gamma_r$, a Dirichlet boundary condition is imposed:

$$u = u_{ex}.$$

2.1. High-order FEM

This section describes the formulation of the high-order, polynomial finite-element method ($p$-FEM) for solving the Helmholtz equation.

2.1.1. Variational formulation

Using the Galerkin method, equations (1), (2) and (3) can be combined into a weak variational statement. It consists in finding the solution $u \in U$ such that:

$$\int_\Omega \left( -k^2 u + \nabla v \cdot \nabla u \right) d\Omega = \int_{\Gamma_r} \overline{\nabla (g - iku)} d\Gamma, \quad \forall v \in V,$$

where $U = \{ u \in H^1(\Omega), u = u_{ex} \text{ on } \Gamma_d, V = \{ v \in H^1(\Omega), v = 0 \text{ on } \Gamma_d \}$ and $\overline{\cdot}$ denotes the complex conjugate of $\cdot$. To define a finite-element approximation of the variational formulation (4), the domain $\Omega$ is partitioned into a set of conforming non-overlapping triangular finite elements $\Omega_e$ of typical size $h$.

2.1.2. High-order shape functions

The approximate solution $u^h$ and the associated test function $v^h$ are constructed using the classical $H^1$-conforming hierarchical polynomial shape functions. This section provides a summary of the definitions of these shape functions. A detailed description can be found in Šolín et al. [6]. The interpolation of the solution is constructed on the reference triangular element defined as $[\xi, \eta] \in \mathbb{R}^2; -1 < \xi, \eta; \xi + \eta < 0$. This reference element is equipped with the following affine coordinates:

$$\lambda_1(\xi, \eta) = \frac{\eta + \frac{1}{2}}{2}, \quad \lambda_2(\xi, \eta) = -\frac{\eta + \xi}{2}, \quad \lambda_3(\xi, \eta) = \frac{\xi + \frac{1}{2}}{2}.$$

The construction of the higher order shape functions rely on the Lobatto polynomials

$$l_0(x) = \frac{1 - x}{2}, \quad l_1(x) = \frac{x + 1}{2}, \quad l_q(x) = \frac{2q - 1}{2} \int_{-1}^{1} L_{q-1}(s) ds, \quad \text{for } -1 \leq x \leq 1,$$

with $L_q$ the Legendre polynomial of order $q \geq 2$. These functions verify the following orthogonality property:

$$\int_{-1}^{1} \frac{\partial l_i}{\partial \xi} \frac{\partial l_j}{\partial \xi} d\xi = 0, \quad \text{for } i \neq j.$$
which ensures an optimal conditioning of the stiffness matrix for Helmholtz problems. It is also necessary to introduce the so-called kernel functions

$$\Phi_q(x) = l_{q+2}(x) / [l_0(x)l_1(x)].$$  \hfill (8)

In each physical triangular element \( \Omega \), the numerical solution \( u^h \) is sought as a sum of contributions from nodal, edge and bubble shape functions. Nodal shape functions are equal to unity at their node of definition and vanish at all other nodes:

$$\phi^{(1)} = \lambda_2, \quad \phi^{(2)} = \lambda_3, \quad \phi^{(3)} = \lambda_1.$$  \hfill (9)

Edge functions are equal to the Lobatto polynomials on the edge they are associated to, and vanish on the other edges of the element:

$$\phi_q^{(e)} = \lambda_2 \lambda_3 \Phi_{q-2}(\lambda_3 - \lambda_2), \quad \phi_q^{(e)} = \lambda_3 \lambda_1 \Phi_{q-2}(\lambda_1 - \lambda_3), \quad \phi_q^{(e)} = \lambda_1 \lambda_2 \Phi_{q-2}(\lambda_2 - \lambda_1),$$  \hfill (10)

with \( 2 \leq q \leq p \). Finally, the bubble shape functions vanish on the boundaries of the element but are non-zero inside the element:

$$\phi_q^{(b)} = \lambda_1 \lambda_2 \lambda_3 \Phi_{q-1}(\lambda_3 - \lambda_2)\Phi_{q-3}(\lambda_2 - \lambda_1),$$  \hfill (11)

with \( q_1, q_2 \geq 1 \) and \( q_1 + q_2 \leq p - 1 \).

On a given triangular element \( \Omega \), the discrete solution of order \( p \) is then expressed as:

$$u^h(\xi, \eta) = \sum_{r=1}^{3} \phi^{(r)}(\xi, \eta)u^{(r)} + \sum_{q=2}^{p} \sum_{r=1}^{3} \phi_q^{(e)}(\xi, \eta)u_q^{(e)} + \sum_{q_1=2}^{p} \sum_{q_2=2}^{p-1} \phi_{q_1, q_2}^{(b)}(\xi, \eta)u_{q_1, q_2}^{(b)},$$  \hfill (12)

where \( u^{(r)} \), \( u_q^{(e)} \) and \( u_{q_1, q_2}^{(b)} \) are the degrees of freedom (DOFs) associated to the nodal, edge and bubble functions respectively. On a given triangular element of order \( p \), there are therefore 3 nodal functions, \( 3(p - 1) \) edge functions and \( (p - 1)(p - 2)/2 \) bubble functions.

An important property of the family of Lobatto shape functions is their hierarchical structure. The shape functions for order \( p \) are retained to construct the set of shape functions at order \( p + 1 \). As a consequence, the orders of the elements can easily be varied across the mesh without requiring any specific treatment to ensure continuity between elements, which is particularly useful for \( p \)-adaptive and \( hp \)-adaptive finite-element methods. In this study, however, the focus is on the performance of the \( p \)-FEM method at fixed order, and \( p \) is considered constant across the whole domain.

For the evaluation of the integrals in the \( p \)-FEM formulation, Gauss–Legendre quadratures are used. For the application cases considered in this paper, all these quadratures are exact, except for the integrand involving \( g \) which may be non-polynomial. For this particular line integral, the order of the numerical quadrature is fixed to \( 2p + 2 \). A large number of tests have also been performed using quadrature orders up to \( 2p + 10 \), without inducing any significant change in the results, thereby indicating that the resulting integration error can be considered negligible.

### 2.1.3. Static condensation

The bubble functions defined in vanish on the element boundaries and are therefore not coupled to the neighbouring elements. These internal degrees of freedom can be eliminated at the element level prior to assembly and do not appear in the global system matrix. This static condensation procedure does not affect the final solution and is a very effective technique to reduce the size and improve the conditioning of the global system matrices. The internal values of the solution are recovered during the post-processing stage by solving small linear systems for each element. The benefits of condensation tend to become more significant as the polynomial order \( p \) increases because the number of bubble functions scales like \( p^2 \) whereas the number of edge functions scales like \( p \). Note however that linear and quadratic triangular elements do not involve bubble functions and cannot benefit from static condensation.

### 2.1.4. Error estimates

It is useful to recall some theoretical results on the performance of high-order finite elements so as to support the discussion of the results. For the \( hp \)-version of the finite element method for smooth solutions of the Helmholtz
equation, Ihlenburg and Babuška [39] have obtained upper bounds for the global relative error in $H^1$ semi-norm

$$E \leq C_1 \left( \frac{kh}{2p} \right)^p + C_2 k \left( \frac{kh}{2p} \right)^{2p},$$

where $C_1$ and $C_2$ are independent of $k$ and $h$ and are weak functions of the order $p$. The first term in this expression represents the interpolation error and has a dependency of order $p$ with $kh$. This error can be controlled by using a sufficient number of elements per wavelength, given by $2\pi/(kh)$. The second term is associated with the dispersion error and the pollution effect. The dispersion error is the difference between the theoretical wavenumber and the wavenumber actually observed in the numerical model. From this error estimate, the advantages of resorting to higher order shape functions appears clearly for smooth solutions. Exponential convergence rates are obtained with $p$-refinement, while convergence with $h$-refinement is only algebraic.

At high resolution (i.e at low values of $kh$), the interpolation error dominates and the $H^1$-norm error converges asymptotically as $O(kh/2p)^p$. In the $L^2$-norm, the actual rate of convergence of the interpolation error is increased by one level and follows $O(kh/2p)^{p+1}$ [40]. These theoretical error estimates are found to match closely with numerical experiments, see for instance [41].

2.2. Wave-based DGM

This section describes the formulation of the wave-based DGM [18] for solving the Helmholtz equation. Only the key aspects of the method that are required for the discussion of the results are described here. The reader is referred to [19] for a more detailed presentation.

2.2.1. Governing equations

The Helmholtz equation [1] is first reformulated as a set of linear, first-order equations of the form:

$$i\omega \mathbf{u} + \frac{\partial \mathbf{A} \mathbf{u}}{\partial x} + \frac{\partial \mathbf{B} \mathbf{u}}{\partial y} = 0,$$

by defining

$$\mathbf{u} = \begin{bmatrix} \rho' \\ \rho_0 u' \\ \rho_0 v' \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ c_0^2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ c_0^2 & 0 & 0 \end{bmatrix},$$

where $\rho_0$ is the mean density of the fluid. The unknowns $\rho'$, $\rho_0 u'$ and $\rho_0 v'$ are the linear perturbations of density, $x$-momentum and $y$-momentum respectively. The acoustic pressure is simply retrieved in the post-processing phase by defining

$$P = c_0^2 \mathbf{u} \cdot \mathbf{n}.$$

2.2.2. Variational Formulation

As for the $p$-FEM presented above, the computational domain $\Omega$ is divided into a collection of elements $\Omega_e$ but unlike $p$-FEM the solution is allowed to be discontinuous across the element edges. After integration by parts in each element the variational formulation for equation (13) reads:

$$\sum_e \int_{\Omega_e} \left( i\omega \mathbf{v} \mathbf{u} - \frac{\partial \mathbf{A} \mathbf{u}}{\partial x} - \frac{\partial \mathbf{B} \mathbf{u}}{\partial y} \right) d\Omega + \sum_{\gamma} \int_{\partial\Omega_e} (\mathbf{v}^T \mathbf{A} \mathbf{n}_e + \mathbf{v}^T \mathbf{B} \mathbf{n}_e) d\Gamma = 0,$$

where $^T$ denotes the conjugate transpose. $\mathbf{v}$ is the test function associated to the solution $\mathbf{u}$ and $\mathbf{n}_e = (n_x, n_y)$ is the outward unit normal to $\partial\Omega_e$.

After introducing the flux matrix $\mathbf{F} = \mathbf{A} \mathbf{n}_e + \mathbf{B} \mathbf{n}_e$ and rearranging terms in the first integral, equation (15) becomes:

$$- \sum_e \int_{\Omega_e} \left( i\omega \mathbf{v} + \mathbf{A}^T \frac{\partial \mathbf{v}}{\partial x} + \mathbf{B}^T \frac{\partial \mathbf{v}}{\partial y} \right)^T \mathbf{u} d\Omega + \int_{\partial\Omega} \mathbf{v}^T \mathbf{F} \mathbf{u} d\Gamma + \sum_{\gamma} \sum_{e' \in e} \int_{\partial\Omega_e} (\mathbf{v}^T \mathbf{F} u_{e'} + (\mathbf{v}^T \mathbf{F}) u_{e'}) d\Gamma = 0,$$
where $\Gamma_{e,e'}$ is the edge between the element $e$ and $e'$. The continuity of the normal flux across the edge between the elements $e$ and $e'$ is directly enforced by writing:

$$\mathbf{F}_e \mathbf{u}_e = -\mathbf{F}_{e'} \mathbf{u}_{e'} = \mathbf{f}_{e,e'}(\mathbf{u}_e, \mathbf{u}_{e'}) ,$$

where $\mathbf{f}_{e,e'}$ is the so-called numerical flux which is discussed below. In the wave-based DGM, the solution in each element is assumed to satisfy the homogeneous equation:

$$i\omega \mathbf{u} + \mathbf{A} \frac{\partial \mathbf{u}}{\partial x} + \mathbf{B} \frac{\partial \mathbf{u}}{\partial y} = \mathbf{0} .$$

(18)

In addition, the test function $\mathbf{v}$ in each element is taken to be a solution of the adjoint equation which can be easily identified from (16):

$$-i\omega \mathbf{v} - \mathbf{A}^T \frac{\partial \mathbf{v}}{\partial x} - \mathbf{B}^T \frac{\partial \mathbf{v}}{\partial y} = \mathbf{0} .$$

(19)

This particular choice of test function is crucial as it allows to remove the element integrals from the formulation and the variational formulation finally reads:

$$\sum_e \sum_{e' < e} \int_{\Gamma_{e,e'}} (\mathbf{v}_e - \mathbf{v}_{e'})^T \mathbf{f}_{e,e'}(\mathbf{u}_e, \mathbf{u}_{e'}) \, d\Gamma + \int_{\partial \Omega} \mathbf{v}^T \mathbf{F} \mathbf{u} \, d\Gamma = \mathbf{0} .$$

(20)

2.2.3. Numerical flux

The efficiency of a discontinuous Galerkin method depends to a large extent on the choice of the numerical flux. The different categories of numerical fluxes available will not be reviewed here and the reader is referred to [42]. In the current formulation, the upwind flux-vector splitting method is used. The flux on the edge between two elements is expressed in terms of characteristic waves in the direction normal to the edge [43]. These characteristic waves can be obtained by diagonalising the flux matrix $\mathbf{F} = \mathbf{W} \Lambda \mathbf{W}^{-1}$. The matrix of eigenvalues $\Lambda$ contains the phase velocity of the characteristic waves and $\mathbf{W}$ is the matrix of eigenvectors. They are used to split the flux matrix $\mathbf{F}$ into the incoming and outgoing flux terms by writing

$$\mathbf{F} = \mathbf{F}^+ + \mathbf{F}^- , \quad \text{with } \mathbf{F}^\pm = \mathbf{W} \Lambda^\pm \mathbf{W}^{-1}$$

where $\Lambda^\pm$ is the diagonal matrix of the positive or negative eigenvalues. For a given edge, the outgoing flux is calculated using the inner solution while the incoming flux is calculated using the outer solution, resulting in the following expression for the numerical flux:

$$\mathbf{f}_{e,e'}(\mathbf{u}_e, \mathbf{u}_{e'}) = \mathbf{F}^+ \mathbf{u}_e + \mathbf{F}^- \mathbf{u}_{e'} .$$

(21)

The variational formulation reads:

$$\sum_e \sum_{e' < e} \int_{\Gamma_{e,e'}} (\mathbf{v}_e - \mathbf{v}_{e'})^T (\mathbf{F}^+ \mathbf{u}_e + \mathbf{F}^- \mathbf{u}_{e'}) \, d\Gamma + \int_{\partial \Omega} \mathbf{v}^T \mathbf{F} \mathbf{u} \, d\Gamma = \mathbf{0} .$$

(22)

In the case of the Helmholtz equation formulated with [13,14] the calculation of $\Lambda$ yields one incoming acoustic wave and one outgoing acoustic wave, with $c_0$ and $-c_0$ as respective phase speeds. The expressions for $\Lambda$ and $\mathbf{W}$ are available in [19] and are not repeated here.

2.2.4. Boundary condition

A variety of boundary conditions can be formulated within the wave-based DGM and a systematic approach for this is presented in [44]. In the present paper, ghost cells and Dirichlet boundary conditions are used.

In the context of finite volume methods and DG methods, the use of ghost cells is a standard technique to generate a prescribed solution within the computational domain [42]. This technique follows naturally from the numerical flux presented above and the idea is to treat the outer boundary of the computational domain as an internal edge between
The variational formulation finally becomes:

\[ \mathbf{F} \mathbf{u} = \mathbf{F}^+ \mathbf{u} + \mathbf{F}^- \mathbf{u} = \mathbf{F}^+ \mathbf{u} + \mathbf{F}^- \mathbf{g} \, . \]

The key difference with an internal edge is that the outer solution is known and will be denoted \( \mathbf{g} \). The incoming flux is therefore expressed in terms of this prescribed solution by writing \( \mathbf{F}^+ \mathbf{u} = \mathbf{F}^- \mathbf{g} \). It should be noted that this approach is equivalent to the Robin boundary condition (2) used for the \( p \)-FEM. The forcing term \( \mathbf{g} \) for the wave-based DGM can in fact be expressed in terms of the target solution \( u_{ex} \) for the \( p \)-FEM:

\[ \mathbf{g} = -\frac{1}{i \omega} \begin{bmatrix} -i \omega u_{ex}/c_0^2 \\ \partial u_{ex}/\partial x \\ \partial u_{ex}/\partial y \end{bmatrix} . \]  

For the Dirichlet boundary condition (3), the method outlined in (44) is used to modify the variational formulation accordingly. Imposing the pressure on the boundary amounts to specifying a relation between the outgoing and incoming characteristic waves. The flux term on the boundary can then be written \( \mathbf{F} \mathbf{u} = \mathbf{F}_d \mathbf{u} + \mathbf{g}_d \) with the following definitions (44):

\[ \mathbf{F}_d = \begin{bmatrix} c_0 & n_x & n_y \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad  \mathbf{g}_d = \begin{bmatrix} -u_{ex}/c_0 \\ u_{ex}n_x \\ u_{ex}n_y \end{bmatrix} . \]  

The variational formulation finally becomes:

\[ \sum_{e} \sum_{e'} \int_{V_{e,e'}} (\mathbf{v}_e - \mathbf{v}_{e'})^T (\mathbf{F}^+ \mathbf{u}_e + \mathbf{F}^- \mathbf{u}_{e'}) \, d\mathbf{r} + \int_{V_e} \mathbf{v}^T \mathbf{F}^+ \mathbf{u} \, d\mathbf{r} + \int_{V_e} \mathbf{v}^T \mathbf{F}_d \mathbf{u} \, d\mathbf{r} - \int_{V_e} \mathbf{v}^T \mathbf{g}_d \, d\mathbf{r} - \int_{V_e} \mathbf{v}^T \mathbf{g}_r \, d\mathbf{r} . \]

2.2.5. Interpolation basis

With Trefftz methods the solution is approximated using local, canonical solutions of the governing equations. Various choices of approximating functions are available: propagating plane waves [35, 14, 17, 27], evanescent plane waves [29, 45, 46], Bessel functions [13], Green’s functions [47]. Some elements of comparisons of these methods can be found in [48, 36, 49, 19, 37].

In the wave-based DGM the approximate solution \( u_e^p \) and test function \( v_e^p \) in each element \( \Omega_e \) are constructed as sums of propagating plane waves:

\[ u_e^p(x) = \sum_{n=1}^{N_e} a_e^p_{n} \exp(-ik_n \mathbf{r} \cdot \mathbf{x}) , \quad v_e^p(x) = \sum_{m=1}^{N_e} b_m^p \exp(-ik_m \mathbf{r} \cdot \mathbf{x}) . \]  

With this approximation the degrees of freedom in each element are the complex wave amplitudes \( a_e^p \) and \( b_m^p \). While the amplitudes \( a_e^p \) and \( b_m^p \) are different, the same sets of wave directions \( \theta_n \) and wavenumbers \( k_n \) are used for the solution and the test function.

The unit vectors \( \mathbf{e}_n \) represent the directions of the plane waves. In the framework of the DGM, the number and the orientations of the plane waves can easily be chosen independently in each element. For simplicity, in this study the number \( N_e \) of plane waves is used in every element and the wave directions are evenly distributed on the interval \([0, 2\pi]\). In the literature, both theoretical results and numerical evidence indicate that \( N_e = 2n \) and \( N_e = 2n - 1 \) (where \( n \in \mathbb{N}^+ \)) provide the same order of convergence, see for instance [50, 21]. This implies that using an odd number of plane waves offers a gain in performance [19, 20]. Therefore, for the sake of conciseness, only odd numbers of plane waves will be shown in this study, with \( N_e \) varying form 3 to 25. While not shown here, results for even numbers of plane waves have also been calculated and the performance and conditioning of the wave-based DGM was indeed found to be slightly below those obtained with odd values of \( N_e \).

The formulation of the wave-based DGM requires that the solution in each element is an exact solution of equation (18). It follows that the wavenumbers \( k_n \) and vector amplitudes \( a_n \) introduced in (25) have to be solution of the following eigenvalue problem:

\[ (\mathbf{A} \cos \theta_n + \mathbf{B} \sin \theta_n) a_n = \frac{i \omega}{k_n} a_n . \]
This equation represents the dispersion relation of the waves propagating in the $\theta_n$ direction.

Similarly, the test function $v^h$ has to satisfy the adjoint equation (19) and it follows that the wavenumbers $k_n$ and vector amplitudes $l_n$ are solution of the adjoint dispersion relation:

\[
(A \cos \theta_n + B \sin \theta_n)^T l_n = \frac{\omega}{k_n} l_n.
\]  (28)

In the case of the Helmholtz equation formulated with (13–14), for a given direction $\theta_n$ one finds two acoustic waves propagating in opposite directions with wavenumbers $\pm \omega/c_0$. The expressions for the vectors $r_n$ and $l_n$ are available in (19) and are not repeated here.

The wavenumbers $k_n$ and the vector amplitudes $r_n$ and $l_n$ are exact solutions of the dispersion relations (27–28). By construction the approximate solution will directly include these properties.

It can be noted that, using this particular wave basis, the variational formulation (25) only involves integrals of exponentials on the edges between elements, which are inexpensive to compute. This is in contrast with the PUFEM, for instance, which involves costly integrals of polynomial-exponential products (35). The integrals on the boundaries $\Gamma_d$ and $\Gamma_i$ involving the source terms are evaluated in closed form whenever possible, or using Gauss–Legendre quadrature rules with a large number of integration points to ensure the results are independent of this numerical quadrature.

2.2.6. Error estimates

Theoretical error estimates for the $h$-version of the UWVF are given in (51), for the Helmholtz equation. It is shown that for smooth solutions, the global discretization error measured in the $L^2$-norm behaves asymptotically like $h^{(N_r-1)/2}$. However numerical tests carried out for different frequencies and number of plane waves indicate that this expression under estimates the actual convergence rates. The convergence rates obtained from numerical experiments reported in (51) are found to be close to $h^{(N_r+1)/2}$, where $[x]$ denotes the integer part of $x$. Similar error estimates are also derived for $h$-refinement by Cessenat et al. in (14), although for a slightly different error norm.

In a series of papers, Hiptmair et al. (20, 22, 23) discussed the behavior of the $p$ and the $hp$-version of the wave-based DGM. The main result that, for a fixed mesh, the $L^2$-norm of the discretization error converges exponentially in the square root of the number of degrees of freedom, including in the presence of strong mesh refinements (22). Concerning the dispersion error, likely to dominate in the lower resolution regime, it is analyzed for both triangular and quadrilateral meshes in 2D in (50). The relative dispersion error is shown to behave asymptotically like $(kh)^{N_r-1}$ for large values of $kh$, hence exhibiting an exponential decrease as the number of plane waves is increased.

3. Description of the test cases

In this section, the test cases used to assess the $p$-FEM and the wave-based DGM are introduced. The performance of high-order methods are known to be problem dependent (12), therefore a total of four different benchmark problems are used to assess the ability of these methods to tackle a large variety of problems. In all computations, all quantities are non-dimensionalised, the speed of sound $c_0$ and the density $\rho_0$ are set to unity. Examples of solutions for each of the four problems are shown in Figure 1.

The first test case is a simple plane wave propagating on a square domain (see Fig. 1(a)). This rudimentary scenario allows for a detailed analysis of the dispersion properties of the numerical methods such as the anisotropy (i.e. the variation in numerical accuracy depending on the angle of propagation). It has been used in previous studies to examine the performance of the wave-based DGM (13), the Partition of Unity method (35) and the Discontinuous Galerkin method with Lagrange multipliers (28).

The computational domain is a square of size $a$ and is represented by an unstructured mesh of triangular elements of typical size $h$ (see Fig. 2(a)). A single propagating plane wave with direction $\theta_0$ is generated inside the domain:

\[u_{ex} = e^{-i \theta_0 x},\]

where $\theta_0 = (\cos \theta_0, \sin \theta_0)$. For the $p$-FEM this is achieved using the Robin condition (2) and for the wave-based DGM the ghost cells are used with (23).
Figure 1: Examples of analytical solutions (real part) of the four benchmark problems ($kL = 50$).

(a) Propagating plane wave ($\theta_0 = 45^\circ$).
(b) Propagating spinning wave ($m = 10$).
(c) Evanescent spinning wave ($m = 19$)
(d) Corner solution.

Figure 2: Examples of unstructured meshes used for the test cases.

(a) Plane wave problem ($h/a = 0.15$).
(b) Spinning wave problem ($h/R = 0.50$).
(c) Corner problem ($h/b = 0.25$).
While providing insight into the accuracy of the numerical schemes, this plane wave solution is not representative of realistic applications where the sound fields typically contain a wide range of wave directions. In addition, realistic problems can also involve evanescent waves in addition to propagating waves, and these can be difficult to capture by some numerical methods \cite{36,13}.

This motivates the introduction of a second test case which is the simulation of a single cylindrical harmonics of order \( m \) which will be referred to as the spinning wave problem. Examples of such solutions are shown in Figures 1(b) and 1(c) and it can be seen that all wave directions are equally represented in this problem. The analytical solution is given in cylindrical coordinates:

\[
\psi_{ex}(r, \theta) = H^{(2)}_{m}(kr)e^{im\theta},
\]

where \( H^{(2)}_{m} \) is the Hankel functions of the second kind of order \( m \) which determines the number of spiral waves around the circumference.

This test case is solved on an annular computational domain with inner and outer radii \( R \) and \( 2R \), respectively. The exact spinning waves \( u_{ex} \) are enforced on the inner and on the outer boundary using the Robin condition \( \psi_{ex} \) for \( p \text{-FEM} \) or the ghost cell formulation with \( u_{ex} \) for the wave-based DGM. An example of unstructured triangular mesh used for this test case is shown in Figure 2(b). In order to avoid introducing any additional source of error in the model, the circular geometry is represented exactly. For \( p \text{-FEM} \), this is achieved by using a linear coordinate mapping between the reference element and the cylindrical coordinates \( r \) and \( \theta \) (instead of Cartesian coordinates \( x \) and \( y \)). For the wave-based method the edges of the elements on the boundaries are defined as arcs rather than straight lines. For both methods, this leads to an exact representation of the cylindrical boundaries of the domain.

It is worth noting that this test case is closely related to the well-known problem of a plane wave scattered by a cylinder since the solution for the latter is a sum of spherical harmonics, see for instance \cite{37,44}. It is preferable to consider different cylindrical harmonics individually. For \( m < kr \) the sound field decays slowly like \( 1/\sqrt{r} \) away from the inner boundary. However, for \( m > kr \) there is a region near the inner boundary where the sound field decays at a much faster rate \( 1/r^m \). These harmonics represent creeping waves that propagate along smooth convex surfaces (see Section 9.5 in \cite{52}). While creeping waves do radiate sound to the far field, they play the same role as evanescent waves radiating from a flat surface. As will be shown below, the large gradients present in the solutions for \( m > kr \) lead to a change in the performance of the numerical methods. For this reason both propagating (Fig. 1(b)) and evanescent (Fig. 1(c)) spinning wave harmonics will be considered and are referred to as the second and third test cases, respectively.

All the benchmark problems described above involve smooth solutions, but the performance of high-order schemes is known to deteriorate significantly when dealing with non-smooth solutions, including the \( p \text{-FEM} \) \cite{53} and the wave-based DGM \cite{36}. Therefore, the fourth test case involves a corner solution. This test case, shown in Figure 1(d) was also used in previous studies \cite{36,9}. The analytical expression for the pressure field is given by

\[
u_{ex}(r, \theta) = J_{2/3}(kr)\sin\left(\frac{2}{3}\theta\right), \quad \text{for } 0 < \theta < \frac{3}{2}\pi,
\]

where \( J_{2/3} \) is the Bessel function of the first kind. While \( u_{ex} \) itself remains bounded, its gradient is singular at the corner. This will allow to illustrate how these methods cope with non-smooth solutions.

As shown in Figure 1(d) an L-shaped computational domain is used for this test case. On the two edges of length \( b \) which meet at the origin, a homogeneous Dirichlet condition is enforced: \( u = 0 \). This is done using the Dirichlet boundary condition \( (3) \) for \( p \text{-FEM} \) and the formulation \( (25) \) with \( (24) \) for wave-based DGM.

For all the other boundaries the Robin condition \( (2) \) is used for \( p \text{-FEM} \) and the ghost cells \( (23) \) are used for the wave-based DGM. Unstructured triangular meshes will be used and an example is shown in Figure 2(c). Local grid refinement near the corner will also be considered to improve the accuracy of the numerical methods. This issue will be addressed in more details when discussing the results for this test case in Sections 5 and 6.

4. Measures of accuracy and costs

For all the test cases described above a number of parameters can be varied: the angular frequency \( \omega \), the element size \( h \) and the approximation order (that is the polynomial order \( p \) for the \( p \text{-FEM} \) and the number of plane waves \( N_w \) for
the wave-based DGM). Since the sizes and shapes of the computational domains vary between the different test cases it is more representative to define the frequency in terms of the Helmholtz number $kL$. The length $L = \sqrt{\text{area}(\Omega)}$ is indicative of the size of the domain and different test cases solved with the same Helmholtz number $kL$ can be expected to be as demanding.

With high-order methods the element size $h$ is not a valid indicator of the resolution of a given numerical method, since the order $p$ or the number of plane waves $N_w$ can be varied independently of $h$. In this work the resolution is measured by calculating the number of degrees of freedom per wavelength, which is defined as follows

$$D_\lambda = \frac{2\pi}{kL} \left( \sqrt{N_{\text{DOF}}} - 1 \right).$$  \hspace{1cm} (31)

This definition is constructed by analogy with a regular Cartesian grid on a square domain of side $L$ and involving $N_{\text{DOF}}$ degrees of freedom (more details are provided in [18]). Unless explicitly stated the mesh resolution will be measured using $D_\lambda$ instead of the element size $h$.

The accuracy of the numerical solutions is assessed by calculating the relative $L^2$ error

$$E_{L^2(\Omega)} = \frac{\|u^h - u_{\text{ex}}\|_{L^2(\Omega)}}{\|u_{\text{ex}}\|_{L^2(\Omega)}}, \quad \text{with } \| \cdot \|_{L^2(\Omega)} = \left( \int_{\Omega} | \cdot |^2 \, d\Omega \right)^{1/2}. \hspace{1cm} (32)$$

The $L^2$ norm is evaluated using high-order Gauss–Legendre quadrature rules on each element, with orders $2p + 10$ for the high-order FEM and $N_w + 10$ for the wave-based DGM method, respectively. The $H^1$ relative error was also measured, but it leads to the same conclusions and will not be reported here for conciseness.

Once assembled, both $p$-FEM and wave-based DGM lead to a sparse global system of equations with complex-valued symmetric coefficients. The latter can be solved either using a direct or an iterative approach.

In this study, only direct solution procedures are considered. The development of robust iterative solvers for Helmholtz problems discretised by the standard FEM is undergoing active research [54]. However, the indefiniteness of the operator tends to hinder their performance at higher frequencies. Furthermore, the matrices arising from a high-order discretisation typically exhibit higher condition numbers than that of the linear FEM, which renders the problem even more challenging (high-order bases designed to address this issue have been proposed [55, 56]). For these reasons, iterative solutions will not be considered here and are left for further investigations. Instead, all systems are solved using the Cholesky factorisation algorithm of the multi-frontal sparse direct solver MUMPS (version 4.10.0) [57].

This operation represents the most expensive part of the solution procedure. The cost of solving the global system is monitored in a number of ways:

- The total number of degrees of freedom $N_{\text{DOF}}$ is a first indicator, but in practice it does not necessarily correlate directly with the cost of a numerical method, see for instance [10].
- The number of non-zero entries $N_{NZ}$ in the sparse global matrix is another measure of cost. In particular the memory required to store the matrix before the factorisation is directly proportional to $N_{NZ}$.
- The factorisation memory, as reported by the solver MUMPS, is a crucial parameter because it is generally the bottleneck in solution procedures based on direct solvers. In real-world applications it is often the amount of memory available that dictates the size of the problems that can be solved on a given platform.
- Although this is more dependent on the platform used for the calculations, the time taken by the factorisation will also be reported.
- Finally, it is important to monitor the condition number of the system (here using the 1-norm) as the accuracy, or even success, of the direct solver can be affected by poor conditioning. In addition it is often reported in the literature that wave-based methods are poorly conditioned and this aspect will be carefully examined in the next sections.

Note that for $p$-FEM, when static condensation is applied, $N_{\text{DOF}}$ and $N_{NZ}$ do not take into account the degrees of freedom corresponding to the internal bubble functions, which are eliminated before the global system is assembled. Unless otherwise stated, the cost of solving the linear system will be reported with static condensation. However, its effects on computational cost and conditioning will also be addressed.
5. Comparison of the interpolation properties

Before trying to solve the test cases previously defined using the two methods, a first analysis is performed to assess the interpolation properties of their respective approximation bases. Independently of the numerical formulations, this section discusses the ability of the continuous polynomial basis on the one hand, and the discontinuous plane-wave basis on the other hand, to approximate the solutions of the four different test cases.

This is done by calculating the best interpolation $u_h^{\text{opt}}$ that minimises the $L^2$ error

$$
\int_{\Omega} |u_{\text{ex}} - u_h^{\text{opt}}|^2 \, d\Omega ,
$$

relative to a known exact solution $u_{\text{ex}}$. The best interpolation is written $u_h^{\text{opt}} = \sum a_n \phi_n$ with $\phi_n$ the shape functions. Finding the degrees of freedom $a_n$ that minimise the $L^2$ error defined above involves a simple least-square problem which amounts to solving the linear system $Xa = y$ with

$$
X_{ij} = \int_{\Omega} \phi_i \phi_j \, d\Omega , \quad y_i = \int_{\Omega} \phi_i u_{\text{ex}} \, d\Omega .
$$

These two integrals are also evaluated using high-order Gauss–Legendre quadrature rules, with orders $2p + 10$ and $N_w + 10$ for the $p$-FEM and the wave-based DGM methods, respectively. The relative error on the optimal interpolation is denoted by

$$
E_{L^2(\Omega)}^{\text{opt}} = \frac{\|u_h^{\text{opt}} - u_{\text{ex}}\|_{L^2(\Omega)}}{\|u_{\text{ex}}\|_{L^2(\Omega)}} .
$$

It is worth noting that the results for the wave-based DGM presented in this section also apply to a number of other wave-based methods: UWVF, least-squares method and wave-based DGM with Lagrange multipliers, since all of these methods rely on the same plane-wave approximation basis defined in Section 2.2.3.

5.1. Anisotropy

The anisotropy of the two different approximating bases is best shown using the test case of the single plane wave by varying the incident wave direction $\theta_0$ from 0 to $2\pi$. The relative $L^2$ error of the optimal interpolation is plotted against $\theta_0$ in Figure 3 for various approximation orders and with element size $h/a = 0.1$ and Helmholtz number $kL = 20$.

As expected, both methods exhibit exponential convergence with respect to the order $p$ or the number of plane waves $N_w$. However, the behaviour of the two bases with respect to the propagation angle $\theta_0$ is very different.

Figure 3: Relative $L^2$ error associated with the best interpolation plotted against the direction of the incident plane wave with $h/a = 0.1$, $kL = 20$. Black solid lines: $p$-FEM, red dashed lines: wave-based DGM. Numbers indicates the polynomial order or the number of plane waves.
The polynomial basis is rather insensitive to the plane wave direction and, for a given order \( p \), only small variations of accuracy are observed when varying \( \theta_0 \). This is due to the use of an unstructured mesh (a structured mesh would result in a more pronounced anisotropy governed by the topology of the mesh [41]).

In contrast, the plane-wave basis is very sensitive to the orientation of the incident plane wave. When it is aligned with one of the waves in the basis, the error drops to zero since the exact solution belongs to the interpolation basis [18]. The error reaches its maximum precisely halfway between two wave directions in the basis. Therefore, for the remainder of the paper, for this particular test case, the incident wave direction \( \theta_0 \) will be taken halfway between two plane waves of the basis. This strong anisotropy of the wave-based DGM also justifies the introduction of the spinning wave test case (Figure 1(a)) where all wave directions are equally represented.

5.2. Convergence properties

The interpolation properties of the two approximation bases are now examined with respect to order and mesh resolution. In Figure 4, the relative error of the best interpolation is plotted against the number of degrees of freedom per wavelength \( D_\lambda \). This is done for each value of \( p \) or \( N_w \) by fixing the frequency \( kL = 50 \) and varying the mesh resolution.

For the single plane wave and the propagating spinning wave test cases, the expected behaviour is observed, with an algebraic convergence with element size \( h \) and an exponential convergence with \( p \) or \( N_w \).
For $p$-FEM, the $L^2$ error scales asymptotically as $D_{d}^{-(p+1)}$, which is in agreement with theoretical [40], and numerical [41] results from the literature (see Section 2.1.4). For the wave-based DGM the error scales like $D_{d}^{-(N_{w}+1)/2}$ where $[x]$ denotes the integer part of $x$, which is consistent with the orders of convergence reported in [14] and [51] (see Section 2.2.6). Note that these are rates of $h$-convergence as they have been identified by varying the element size while keeping the order or the number of plane waves constant.

For the wave-based method the interpolation error is not able to go beyond $10^{-5}$. This is due to the poor conditioning of the system solved to obtain the best interpolation. The condition number of the matrix $X$ systematically exceeds $10^{12}$ when the relative $L^2$ error of the best interpolation is below $10^{-4} \%$.

The overall conclusion from Figures 4(a) and 4(b) is that, for smooth solutions, there is no significant gap between the two bases in terms of interpolation error. Using reasonably high polynomial orders or numbers of plane waves, it is possible to achieve similar levels of accuracy with the same number of degrees of freedom per wavelength.

For instance for the plane wave test case, the polynomial basis with order $p = 10$ matches closely the accuracy of the plane wave basis approximation with $N_{w} = 21$.

A similar comparison has been performed for the spinning wave test case and the results are consistent with those reported here. Namely, the accuracy for $p$-FEM is very similar to figure 4(b) and for the wave-based DGM the numerical error is found to be concentrated close to the cylinder which is consistent with the evanescent part of the plane wave basis approximation.

As explained above the solution to the well-known test case of a plane wave scattered by a cylinder can be expressed as a combination of propagating and evanescent cylindrical harmonics. Although not reported here for the sake of brevity, a similar comparison has been performed for this test case and the results are consistent with those reported here. Namely, for smooth solutions, there is no significant gap between the two bases in terms of interpolation error. Using reasonably high polynomial orders or numbers of plane waves, it is possible to achieve similar levels of accuracy with the same number of degrees of freedom per wavelength.

Finally figure 4(d) shows the results for the corner solution obtained with a uniform mesh. The exponential convergence with $p$ and $N_{w}$ is lost for both methods. The interpolation error still decreases but at the same slow rate for all orders. This behaviour is expected when dealing with solutions that are not sufficiently regular (in this case the derivative of the solution is singular), see for instance [12,40]. The use of local $h$-refinement is a valid option to improve both the FEM and the wave-based DGM solutions and will be used in the next section.

For problems with singular solutions or evanescent waves, Trefftz methods can benefit from the introduction of additional $ad$ hoc functions in their approximation basis. The use of evanescent plane waves in the basis is for instance examined to represent the transmission and reflection of sound at an interface between two media in [46] for the UWVF and in [25,29] for the DEM. However, this approach obviously requires a detailed knowledge of the solution behaviour before the numerical method is even formulated. Such practical considerations will be discussed in more detail in Section 7.

6. Performance of the numerical methods

In this section, the efficiency of the two numerical methods is compared. The convergence with respect to $p$- and $h$-refinements is firstly considered, followed by the study of the conditioning properties of the underlying system of equations. Finally, the computational cost at fixed accuracy is assessed.

6.1. Convergence

Figure 5 shows the relative $L^2$ error defined in (32) plotted against the number of degrees of freedom per wavelength for the four benchmark problems at $kL = 50$.

For the plane wave and the spinning wave test cases, these convergence results follow, at least qualitatively, the results in Figure 4 for the best interpolation error. Algebraic and exponential convergence rates are also observed for $h$-refinement and $p$-refinement, respectively. For the wave-based method, it is not possible to reach levels of error below $10^{-6}$ due to poor conditioning (this is discussed in the next Section). Again, the overall conclusion is
Figure 5: Relative $L^2$ error (%) against $D_{\lambda}$ at $kL = 50$; red dashed lines: wave-based DGM; black solid lines: $p$-FEM with condensation. For each curve, $D_{\lambda}$ is increased by refining the mesh while $p$ or $N_w$ is fixed.
that the two methods are able to achieve similar levels of accuracy with the same number of degrees of freedom per wavelength.

In order to provide a more quantitative comparison with the best interpolation error, it is useful to calculate the optimality coefficient $C_{opt} = \frac{E_{L^2(\Omega)}}{E_{opt,L^2(\Omega)}}$ which is the ratio between the numerical solution accuracy $E_{L^2(\Omega)}$ and the best interpolation error $E_{opt,L^2(\Omega)}$. This parameter indicates the efficiency with which the numerical formulation can exploit the interpolation properties of the underlying approximation basis.

An example for the propagating spinning wave test case is provided in Figure 6. A large increase in $C_{opt}$ is observed for both methods in the pre-asymptotic regime, corresponding to low values of $D_\lambda$. This increase can be attributed to the pollution effect, corresponding to the presence of large dispersion errors in the model, which typically dominate for poor resolutions. High-order methods allow to mitigate the pollution effect and this is indeed observed in Figure 6. Increasing the order $p$ or the number of plane waves $N_w$ leads to a significant improvement in the optimality of the methods (i.e. a reduction of $C_{opt}$). It can be noted that, overall, $p$-FEM exhibits slightly better optimality coefficients. In the asymptotic regime where the interpolation error dominates (large $D_\lambda$), the optimality coefficient for $p$-FEM varies between 1.2 and 2 whereas it varies between 2 and 10 for the wave-based DGM. The difference is particularly visible for low orders. For $N_w = 3$ and 5, the DGM formulation is far from optimal.

The results obtained for the evanescent spinning wave and the corner solution, depicted respectively in Figures 5(c) and 5(d) follow qualitatively those obtained for the optimal $L^2$ interpolation error given in Figure 4. For the corner solution the loss of exponential convergence with $p$ or $N_w$ is again clearly visible. An additional observation is that for a relative error below 1% the same convergence rate is observed for all orders $p$. It is interesting to note that for this test case the optimality coefficient is much larger than for the other test cases ($C_{opt}$ varies between 5 and 25). This indicates that not only is the corner solution an issue in terms of interpolation but the expected convergence properties of high-order methods are also lost. A summary of these aspects can be found in [48, Section 5.3].

As illustrated for the same corner solution in [48, 36] for wave-based methods and in [9] for $p$-FEM, $h$-refinement is required to obtain accurate solutions. Results obtained with local mesh refinement are now presented to compare its impact on the accuracy of the two methods. To this end an element size $h_{max}$ is enforced along the boundaries $x = \pm 1$ and $y = \pm 1$. The element size at the corner $x = y = 0$ is given by $h_{max}/C_r$ where the refinement factor $C_r$ is varied from 1 to 400. Three examples of locally refined meshes are given in Figure 7.

In Figure 8 the convergence of the solution is plotted against the value of $C_r$ for a fixed mesh size $h_{max}$ and a fixed frequency ($kL = 50$). Two different regimes can be observed. For moderate refinements (i.e. low values of $C_r$) the

Figure 6: Optimality coefficient against $D_\lambda$ for the propagating spinning wave with $kL = 50$; red lines: wave-based DGM; black lines: $p$-FEM with condensation. For each curve, $D_\lambda$ is increased by refining the mesh while $p$ or $N_w$ is fixed.
error due to the singularity dominates the overall numerical error and hence refining allows to improve the relative $L^2$ error. It is worth noting that the convergence rate with $C_r$ in this regime is nearly independent of the approximation order $p$ or $N_w$ and is similar for both methods. For large values of $C_r$, a plateau is reached and the error stagnates because the overall numerical error is dominated by the approximation of the wave propagation in the rest of the domain.

Figure 8 also shows the impact of mesh refinement on the condition number. For $p$-FEM, modifying $C_r$ has little effect on the conditioning. In contrast, it has a large impact on the behaviour of the wave-based DGM. The introduction of small elements in the neighbourhood of the corner and a large number of plane waves in the basis leads to a rapid growth of the condition number \cite{14, 58}. As a consequence, the number of plane waves cannot be arbitrarily increased and in this example values above $N_w = 15$ could not be considered.

These results illustrate that $h$-refinement has a significantly different impact on the behaviour of the two methods. While it consistently improves $p$-FEM results when the singularity error dominates, it should be used with caution for the wave-based DGM, where it can lead to stability issues. Two options can be considered to further modify the wave-based method for this situation. First, locally adapting the number of plane waves to the element size would control the conditioning by reducing the number of plane waves in small elements. Secondly, the introduction of local ad hoc functions would also be appropriate, for instance by combining plane waves with Bessel functions \cite{13}.
6.2. Conditioning

In Figure 9, the condition numbers are plotted against \( D_3 \) for the propagating spinning wave test case. The results obtained for the plane wave test case are very similar and are not repeated here.

For well-resolved models (i.e., large values of \( D_3 \)), the conditioning of \( p \)-FEM tends to increase slowly with \( D_3 \). In this regime, the polynomial order has a weak effect on the conditioning. Conversely, for poorly resolved models (typically \( D_3 < 10 \)) there is a visible effect of the order, with a progressive deterioration of the conditioning when \( p \) increases. Also shown in Figure 9 is the conditioning of the \( p \)-FEM model without static condensation. In this case, the global system includes the degrees of freedom corresponding to the bubble shape functions in each element. The impact of static condensation on the conditioning is very significant. Without static condensation, the conditioning of the \( p \)-FEM model deteriorates very rapidly when the polynomial order is increased. For this test case, each increment of \( p \) results in a growth of one order of magnitude for the condition number. For instance for \( p = 10 \), there is a difference of 5 or 6 orders of magnitude between the models with and without static condensation. The effect of the resolution rate \( D_3 \) remains comparatively limited. This explains why the use of static condensation is generally recommended to improve the performance of direct \([59]\) as well as iterative \([60]\) solving procedures.

In comparison with \( p \)-FEM, the conditioning of the wave-based method varies more significantly with the mesh resolution and the order \( N_w \) \([14, 58]\). In particular, for a large number of plane waves, the condition number grows very steeply with the mesh resolution. When plotting the relative error against the condition number (see Fig. 9), the two appear to be closely related. Independently of \( N_w \), the results collapse on a single curve indicating that the condition number scales like \( E_{L^2}^{-2} \). Hence, for reasonable levels of accuracy, say for instance 1\%, the conditioning of the wave-based method is not problematic, and is in fact equivalent to the \( p \)-FEM with static condensation. The conditioning of the wave-based method only becomes an issue when targeting very high levels of accuracy.

For both methods different techniques can be used to improve the conditioning, in particular by selecting other approximation bases. For the wave-based method, the use of Bessel function improve the conditioning \([13]\), while Bernstein polynomials have a similar impact on \( p \)-FEM \([4]\).

6.3. Comparison of computational costs

Given the large number of parameters involved, it is convenient to compare the two methods for a fixed accuracy and frequency. In this way, one can assess the requirements needed for the two methods to solve for a particular problem with a given error threshold. In a second stage, we will discuss to which extent the conclusions drawn are applicable to other frequencies and other levels of accuracy. For a given test case and for every order \( p \) or number of plane waves \( N_w \), the mesh size \( h \) is varied until the numerical error reaches the target precision, then the cost and conditioning of this model are recorded. The test case of the spinning wave is used here but the conclusions are equally applicable to the plane-wave test case.
6.3.1. Fixed accuracy and frequency

The frequency is fixed to \( kL = 50 \) and a target accuracy of 1% is considered, which is regularly suggested as relevant for industrial applications.

Figure 10 shows the link between the mesh size and the factorisation time. The timing was measured on an Intel Xeon E5-2670 processor (2.6 GHz) using a single core. Every data point in Figure 10 achieves the same relative error of 1% but the mesh resolution required to do so varies with \( p \) or \( N_w \). As one would expect, low-order models (e.g. \( p = 2 \) or \( N_w = 5 \)) require very fine meshes to meet the target accuracy and result in longer factorisation times. When using higher-order models it is possible to maintain the same error level with larger elements, and the factorisation time is also reduced significantly. The benefit of increasing the polynomial order or the number of plane waves tends to level off for large values of \( p \) and \( N_w \).

Also shown in Figure 10 is the reduction in factorisation time obtained from static condensation (the mesh resolution remains the same with and without static condensation). This benefit is however balanced by the fact that, when using static condensation, the degrees of freedom associated with the bubble functions have to be calculated after the factorisation. From experience, the use of static condensation does not necessarily reduce the overall runtime of the solution procedure, unless the recovery of the internal degrees of freedom in each element is parallelized.

Figure 11 shows the evolution of the condition number and the factorisation memory when \( p \) and \( N_w \) are varied (but the element size \( h \) is adjusted to maintain the error at 1%).

For the propagating spinning wave, see Figure 10(a) increasing \( N_w \) for the wave-based method leads to a consistent reduction in the memory required to solve the problem but the conditioning does not vary significantly. This is consistent with the results in Figure 9 showing that for a fixed accuracy the conditioning is only weakly influenced by the number of plane waves. For \( p \)-FEM with static condensation, increasing the polynomial order also reduces the factorisation memory and induces a moderate increase in conditioning, although the condition number remains comparable to that of the wave-based method. It can be seen that the two methods are able to achieve the same level of accuracy with similar cost and conditioning. Without static condensation, the cost of \( p \)-FEM does not greatly decrease with \( p \) (except between \( p = 2 \) and 3) and the conditioning is significantly higher.

For the evanescent spinning wave, Figure 10(b) the results for \( p \)-FEM are very similar to those for the propagating wave. However for the wave-based method, the condition number and the memory requirements have both increased. The benefit of using more plane waves therefore tends to level off. Again, this is due to the strong gradient in the solution close to the inner boundary of the computational domain.

The behaviour of the factorisation memory can be related to the properties of the sparse, global system matrices generated by the two methods. Figure 12 shows the evolution of the number of non-zero entries (\( N_{NZ} \)) against the number of degrees of freedom (\( N_{DOF} \)). Note that the ratio \( N_{NZ}/N_{DOF} \) provides an indication of the average number of
terms in each row of the matrix. For $p$-FEM with static condensation, increasing the polynomial order leads to a rapid reduction in the number of degrees of freedom. The same applies to the wave-based method when increasing $N_w$. For both methods the sparsity of the global matrices is reduced when increasing $p$ or $N_w$, but the wave-based method generates denser matrices (i.e. with more non-zero entries). This difference is further accentuated when considering the evanescent spinning wave.

### 6.3.2. Influence of the target accuracy

So far, a fixed accuracy of 1% on the relative $L^2$ error was considered, but it is also interesting to examine other levels of target accuracy. Figure 13 shows the factorisation memory and the conditioning as a function of the number of degrees of freedom per wavelength $D_\lambda$. The target error is varied from 10% to $10^{-4}$%. The test case of the propagating spinning wave is used, but the results for the plane wave and the evanescent spinning wave test cases are similar.

For the factorisation memory, the two methods follow the same trends: the memory increases when the target error level is reduced and when $D_\lambda$ increases. The fact that the wave-based method requires slightly more memory than $p$-FEM is observed at all error levels, confirming that this observation is not specific to the 1% target accuracy used above.

The conditioning of the $p$-FEM model tends to increase slightly when the target error level is lowered, mostly for lower resolutions (i.e. $D_\lambda < 10$). However, for the wave-based method, the condition number increases very rapidly to reach levels that are problematic. Again, this is consistent with the results in Figure 9 which shows that it is only for very accurate solutions that the conditioning of the wave-based method becomes a real issue.

### 6.3.3. Influence of the frequency

Finally, it is important to assess whether the trends described above are also observed for other frequencies. To this end, Figure 14 shows the factorisation memory and the conditioning for Helmholtz numbers ranging from $kL = 25$ to 200 for the propagative spinning wave test case (the azimuthal order $m$ is adjusted at each frequency to ensure the wave is propagating). A target error level of 1% is used.

The computational cost is found to scale like $k^2$, which is the expected behaviour for two-dimensional problems. As the polynomial order or the number of plane waves is increased, the two methods are able to achieve 1% error with less memory. For every frequency, the wave-based method requires a larger amount of memory than $p$-FEM, although the difference is relatively small for the lowest frequency ($kL = 25$).

The evolution of the condition number with respect to the frequency is more complex and does not follow a simple trend. For $D_\lambda > 5$, $p$-FEM appears to be better conditioned, but the opposite is generally true for $D_\lambda < 5$. 

![Figure 11: Factorisation memory against condition number for a numerical error of 1% at $kL = 50$. Red dashed lines: wave-based DGM; Blue dotted lines: $p$-FEM without condensation; Black solid lines: $p$-FEM with condensation.](image-url)
7. Practical considerations

So far, the discussion has focused on the intrinsic computational performance of the numerical schemes. To offer a broader perspective on the use of these two methods, it is also useful to consider more practical aspects.

The performance of the wave-based method, and other Trefftz methods in general, is obviously directly linked to the selection of canonical solutions used to construct the basis, and this has a number of practical consequences:

- As already mentioned above, when dealing with evanescent waves, wave-based methods require the introduction of ad hoc interpolating functions directly associated with these types of solutions. On the other-hand, the performance of $p$-FEM is not significantly altered by the presence of evanescent waves.
- Both methods lose $p$-convergence when dealing with singular solutions. For $p$-FEM, local $h$-refinement is sufficient to reach a good level of accuracy, but due to conditioning issues, the use of special functions should be preferred for wave-based DGM [13].
- For problems involving non-uniform coefficients, which have not been considered here, individual plane waves are not solutions of the governing equations. Defining a basis of simple local solutions for this class of problems is not trivial but the recent work of Imbert-Gérard et al provides a way forward [61, 62]. On the other hand, $p$-FEM can accommodate non-uniform coefficients without additional treatment.
The presence of source terms in the domain also requires special techniques to address the fact that the plane waves are solutions of the homogeneous equations. The $p$-FEM can tackle such problems ‘out-of-the-box’ and a number of techniques have been proposed for wave-based methods [14, 63, 18]. It is therefore needed to adjust the Trefftz interpolation basis locally, on a case-by-case basis, which requires a detailed a priori knowledge of the solution. While this is acceptable for applications to academic test cases, it is difficult to envisage for real-world applications where robust, fully automated solvers are required.

In typical applications, it is often needed to repeat a simulation over a range of frequencies. In such situations, the wave-based method requires calculating and assembling the element matrices again for each frequency due to the dependence of the shape functions on frequency, which represents a significant overhead. By contrast, with $p$-FEM the use of hierarchic shape functions allows to store and reuse the element matrices for other frequencies [10].

8. Conclusions

A detailed comparison of a high-order polynomial method ($p$-FEM) and a wave-based method (DGM) has been presented in terms of interpolation properties, performance and conditioning. The main objective was to put in perspective the relative merits of these two categories of discretisations. Given the large number of parameters available in these high-order methods, and the different metrics of performance that can be considered, this paper also aimed at providing an example of methodology to conduct a systematic and thorough comparison of high-order schemes for frequency-domain analysis.

The general argument for the use of wave-based methods is that plane waves are canonical solutions of the underlying equations and are therefore better suited than polynomials to construct an approximation basis. They incorporate key properties of the exact solution such as the wavenumber (and polarisation for electromagnetism for instance) and the expectation is that they provide a more accurate and efficient family of functions to interpolate the solution. The results presented here indicate that this intuitive expectation does not necessarily translate into a clear benefit in terms of interpolation accuracy (for a fixed resolution $D_\lambda$) or performance (for a fixed numerical error). Rather surprisingly, the high-order polynomial method is able to achieve comparable, and in some cases superior, results compared to a wave-based method for solving two-dimensional Helmholtz problems. This observation was found to hold for a range of frequencies and error levels. It should be noted that the use of static condensation is crucial to obtain good performance and conditioning from $p$-FEM. The conditioning of the wave-based method was found to be acceptable and comparable to that of $p$-FEM for an error level of 1%. It is only when very high levels of accuracy are targeted that the conditioning deteriorates significantly.

The general outcome of this comparison is that there is still significant benefits in using polynomial methods, both in terms of performance and robustness, and that more work is required on wave-based methods to fully realize their potential.
It should be noted that the present work used the wave-based DGM as an example of Trefftz method and other wave-based methods should also be considered. The wave-based DGM is closely related to a number of other methods, in particular the UWVF \cite{O. Cessenat, B. Despr`es, Application of an ultra weak variational formulation of elliptic PDEs to the two-dimensional Helmholtz problem, Journal of Computational Physics 182 (1) (2003) 227–238. doi:10.1016/S0021-9991(03)00191-2} and the least-squares method \cite{P. Monk, D. Wang, A least-squares method for the Helmholtz equation, Computer Methods in Applied Mechanics and Engineering 175 (1-2) (2000) 227–238. doi:10.1016/S0045-7825(98)00326-0}. Both methods can be recovered in the DG framework by using different numerical fluxes. Existing results indicate that the change of numerical flux is unlikely to change the performance of the methods sufficiently to alter the conclusions in this paper \cite{G. Gabard, P. Gamallo, T. Huttunen, A comparison of wave-based discontinuous Galerkin, ultra-weak and least-square methods for wave problems, International Journal for Numerical Methods in Engineering 85 (3) (2011) 380–402. doi:10.1002/nme.2979}. But other categories of wave-based methods, such as DEM and PUFEM remain to be compared systematically with high-order polynomial methods. Extending these comparisons to three-dimensional problems would also be useful.

Acknowledgements

This work was performed as part of the CRANE project (Community and Ramp Aircraft NoiSE, www.craneeid.eu) funded by the European Union under the Framework Programme 7. The authors acknowledge the use of the IRIDIS High Performance Computing Facility, and associated support services at the University of Southampton, in the completion of this work.

References


23


