The Fourier version of the Variational Theory of Complex Rays for medium-frequency acoustics

Louis Kovalevsky 1, Pierre Ladevèze *,1,2, Hervé Riou 1

LMT Cachan, Ens Cachan/UPMC/CNRS/PRES UniverSud Paris, 61 Avenue du président Wilson, 94230 Cachan, France

The Variational Theory of Complex Rays (VTCR) is a wave-based computational approach dedicated to the resolution of medium-frequency problems. It uses a variational formulation of the problem which enables one to use any type of shape function within the substructures provided that it satisfies the governing equation. Thus, the solution can be approximated using plane waves, which is very interesting in the medium-frequency vibration domain and also leads to a strong convergence of the method. In the previous works, this was shown in the case of acoustic problems in which the amplitudes of the plane waves were calculated as wavebands. In this paper, we propose a new approximation of these amplitudes based on Fourier series. We show that this approach increases the robustness of the method, makes it more efficient numerically and extends its applicability to somewhat higher frequencies.

1. Introduction

In recent years, the use of numerical simulation techniques in design, analysis and optimization has become an indispensable part of the industrial design process. The standard Galerkin Finite Element Method (FEM [1]) is a well-established computer-aided engineering tool which is commonly used for the analysis of time-harmonic dynamic problems. However, the use of continuous, piecewise polynomial shape functions leads to very large numerical models and, in practice, restricts the application of this prediction technique to the low-frequency range. Furthermore, since the shape functions are not exact solutions of the governing differential equations, a refined discretization is necessary in order to eliminate the associated pollution error [2].

A variety of techniques have been proposed in order to eliminate the problems associated with lower-order finite elements. These are modified versions of the basic finite element method and include, for example, higher-order methods [3], Galerkin least squares method [4–6], predefined reduced bases [7], the quasi-stabilized finite element method [8], the Partition of Unity Method (PUM) [9], the generalized finite element method [10] the Discontinuous Petrov Galerkin method [19] and the quasi optimal Petrov–Galerkin method [11]. The PUM enables one to include information about the solution a priori into the approximation subspace. Compared to standard finite elements, this approach has been shown to decrease the computational complexity considerably [12]. All of these methods have been shown to reduce computation costs, but their domain of application does not cover the whole medium- and high-frequency range.

Besides the FEM and its variations mentioned previously, there is another family of methods, known as Trefftz methods [13], which differs from the FEM by the choice of the shape functions. Instead of using approximate functions, the expansion of the field variables is achieved using exact solutions of the governing differential equations. A first study of Trefftz methods in the context of vibrations was carried out in [14]. Then, the Variational Theory of Complex Rays (VTCR), which is the subject of this paper and is based on the works presented in [15], was introduced in [16]. A special use of the PUM was proposed in [17] and its generalization to Helmholtz problems was presented in [18]. Subsequently, the Wave-Based Method (WBM, [20]), the Ultra Weak Variational Formulation (UWVF, [21,22]) and the element-free Galerkin method ([23]) were introduced. Then, the least-squares method was published in [24]. Finally, the Discontinuous Enrichment Method (DEM, [25]) and the wave boundary element method ([26]) were introduced.

The decisive advantage which is common to all Trefftz methods is that no refined discretization is necessary since exact solutions of the governing equation are used. Therefore, the size of the model and the associated computational effort are considerably reduced.
compared to finite-element-based methods. The main differences among all these approaches are the treatment of the transmission conditions (integration within a Galerkin procedure, minimization of a least-square expression, use of Lagrange multipliers or development of a dedicated variational formulation), and the types of shape functions being used.

The VTCR, which is the subject of this paper, was proposed in [16,27] for the resolution of vibration problems in the medium-frequency range. This formulation is based on a variational formulation of the problem which was developed in order to enable the approximations within the subdomains to be a priori independent of one another and which takes into account both the boundary conditions and the continuity across the interfaces, thus eliminating the need for a specific treatment to guarantee interelement continuity. Any type of shape function can be used within each subdomain, which brings this approach great flexibility and robustness. For vibration problems, the VTCR approximates the solution through a set of wave functions, called complex rays, which satisfy the governing differential equations exactly. Therefore, no residual error affects the governing partial differential equation inside the problem’s domain. However, the functions may violate the boundary conditions. Forcing these residual boundary errors to be equal to zero thanks to the previous variational formulation leads to a small matrix equation whose solution determines the angular distribution of the waves in space. In [28], the VTCR was used to predict the vibrational response of a 3D plate assembly using interior, edge and corner wave functions. In [29], plates with heterogeneities were taken into account. In [30], this theory was extended to shell structures. The calculation of the vibrational response over a range of frequencies was presented in [31]. The use of the VTCR for transient dynamic problems was covered in [32]. Finally, the extension to acoustic problems was presented in [33]. This approach has been shown through many examples to be capable of producing an accurate solution by using only a small number of degrees of freedom (DOFs). In all the works mentioned before, the VTCR used waveband shape functions. The objective of this paper is to study an alternative choice involving a new type of shape function, based on a Fourier series expansion of the amplitudes of the plane waves which propagate within an acoustic cavity.

Following this introduction, Section 2 briefly addresses the reference problem to be solved, which is a steady-state dynamic problem governed by a system of Helmholtz partial differential equations, and describes the VTCR formulation and approximations for two-dimensional bounded Helmholtz problems. Then, Fourier series are introduced in order to discretize the amplitudes of the plane waves. Some general remarks are made, the implementation aspects of this new approach are discussed, and comparative illustrations with the waveband discretization are given. Section 3 focuses on the consequence of the use of the Fourier series discretization on the efficiency of the VTCR. We show that the robustness of the method (in terms of numerical stability) is improved and that the computations are more efficient. Different comparisons with the waveband VTCR and the FEM are presented using examples of diverse complexity. Finally, Section 4 concludes the paper with general remarks and a list of problems which remain to be investigated.

2. Theoretical framework of the VTCR for Helmholtz problems

2.1. The reference problem

Let us consider a general two-dimensional interior dynamic problem defined in an acoustic cavity $\Omega$ filled with a fluid characterized by its sound velocity $c_0$ and its density $\rho_0$. We are interested in the steady-state dynamic behavior of $\Omega$ at a fixed circular frequency $\omega$. All the descriptive quantities can be expressed as complex numbers: an amplitude $Q(x)$ is associated with the quantity $Q(x)e^{i\omega t}$, where $x$ represents the position and $i = \sqrt{-1}$ is the imaginary unit. The problem to be solved is: find the pressure $p \in H^1(\Omega)$ such that

$$\begin{align*}
\Delta p + k^2 p &= 0 \quad \text{in } \Omega, \\
p - Z L_{el}(p) &= h_d \quad \text{over } \partial_2 \Omega, \\
p &= p_d \quad \text{over } \partial_1 \Omega, \\
L_{el}(p) &= \nu_d \quad \text{over } \partial_1 \Omega,
\end{align*}$$

(1)

where $k = \frac{\omega}{c_0}$ is the wave number, $Z$ an impedance coefficient and $L_{el}(\square)$ an operator defined as $L_{el}(\square) = \frac{1}{\rho_m} \frac{\partial}{\partial n} = \frac{1}{\rho_m} n \cdot V(\square)$, $n$ being the outward normal to $\partial_2$, $h_d$ and $\nu_d$ denote respectively a prescribed excitation over $\partial_2 \Omega$, a prescribed pressure over $\partial_1 \Omega$ and a prescribed velocity over $\partial_1 \Omega$. In this paper, it is assumed that Problem (1) has a solution. The uniqueness of the solution is ensured through absorption coefficients introduced in the imaginary part of $k = k_0(1 - i\eta)$ or in the real part of $Z$. These coefficients, whose values for all frequency regimes can be given by specialists of vibrations, ensure that the solution is finite even if the frequency corresponds to an eigenvalue of the problem.

2.2. Variational formulation of the reference problem

Let us consider a partition of cavity $\Omega$ into $n_d$ non-overlapping subcavities $\Omega_e$ (see Fig. 1). Let $\Gamma_1$ and $\Gamma_{E_F}$ denote $\partial_2 \Omega$ and $\partial_1 \Omega$, $\partial_1 \Omega_e$, respectively. Reference problem (1) becomes:

Find the pressure $(p_1, \ldots, p_{n_d}, p_\Omega) \in H^1(\Omega_1) \times \cdots \times H^1(\Omega_{n_d})$ such that:

$$\begin{align*}
\Delta p_e + k^2 p_e &= f \quad \text{in } \Omega_e, \\
p_e - Z_e L_{el}(p_e) &= h_{de} \quad \text{over } \partial_2 \Omega_e, \\
p_e &= p_{de} \quad \text{over } \partial_1 \Omega_e, \\
L_{el}(p_e) &= \nu_{de} \quad \text{over } \partial_1 \Omega_e, \\
L_{el}(p_\Omega) &= \nu_{de} \quad \text{over } \Gamma_{E_F}.
\end{align*}$$

(2)

The terms $p_e, Z_e, L_{el}, h_{de}$ and $\nu_{de}$ correspond respectively to the pressure, the impedance coefficient and the prescribed excitations over cavity $\Omega_e$. The last two equations represent the continuity conditions over $\Gamma_{E_F}$.

The VTCR formulation is obtained by rewriting the boundary value problem (2) in weak form. In order to do that, we introduce the spaces $S_{de}^{\Omega_e}$ and $S_{de}^{\Omega_\Omega}$ of the functions which satisfy respectively the homogeneous and nonhomogeneous Helmholtz equations (i.e. Eq. (1a) with $f = 0$ and $f \neq 0$) in each subcavity $\Omega_e$.

![Fig. 1. A general bounded 2D Helmholtz problem defined in $\Omega$ and partitioned into $n_d$ non-overlapping subcavities $\Omega_e$.](image-url)
\( S_{ad,0} = \{ p_e \in H^1(\Omega_e)| \Delta p_e + k^2 p_e = 0 \text{ in } \Omega_e \} \), \( S_{ad} = S_{ad,0} + p_E^k \quad (3) \)

where \( p_E^k \) is a particular solution of \( \Delta p_e + k^2 p_e = f \) (e.g., \( p_E^k(x) = \int_{\Omega} f(x)G(x,y)ds(y), G(x,y) \) being the Green function of the Helmholtz equation).

The VTCR formulation is:

Find \( p \in S_{ad} \) such that

\[
\text{Re} \left\{ \sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} \left( (1-Z_{E})(p_e) - h_{ad} \right) \cdot \nabla (\partial p_e) \, d\Gamma \right. \\
+ \sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} \nabla (p_e - p_{ad}) \cdot \nabla (\partial p_e) \, d\Gamma + \sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} \nabla (p_e - v_{ad} \cdot \partial p_e) \, d\Gamma \\
+ \sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} \left( p_e \cdot \nabla (p_e - p_{ad}) - \nabla \cdot (p_e - p_{ad}) + \nabla (p_e + p_{ad}) \cdot \nabla (p_e + p_{ad}) \right) \, d\Gamma \left\} = 0 \forall (\partial p_e \in S_{ad}) , \quad (4)
\]

where \( \Box \) and \( \text{Re} (\Box) \) represent respectively the complex conjugate and the real part of the complex quantity \( \Box \).

In practice, one sets \( p_e = p_E^k + p_P^k \) in Eq. (4) and recasts it in terms of the unknowns \( p_P^k \), with the known particular solution \( p_E^k \) contributing to the right-hand side. This modification will be assumed implicitly throughout the rest of the paper and we will focus on the resolution of the VTCR formulation for \( p_P^k \) (\( 1 \leq E \leq N_e \)).

The equivalence between (2) and (4) can be proven by assuming some regularity and the existence of a solution of the PDE in its strong form. Obviously, this solution is also a solution of (4). Therefore, if it can be proven that (4) has no other solution, then (2) and (4) are equivalent.

Let us consider two solutions of (4) and denote \( \tilde{p} \) the difference between the two. Because of (4), we have

\[
\text{Re} \left\{ \sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} \left( (1-Z_{E})(p_e) - \nabla \cdot (\partial p_e) \right) \, d\Gamma \right. \\
+ \sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} \tilde{p}_e \cdot \nabla (\partial p_e) \, d\Gamma + \sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} \nabla \cdot (\tilde{p}_e) \cdot \partial p_e \, d\Gamma + \sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} \left( \tilde{p}_e \cdot \nabla \cdot (\partial p_e) - \nabla \cdot (\partial \tilde{p}_e) + \nabla \cdot (\tilde{p}_e + \partial \tilde{p}_e) \right) \, d\Gamma \left\} = 0, \forall (\partial p_e \in S_{ad}) . \quad (5)
\]

Taking \( \partial \tilde{p} = \hat{p}, (5) \) reduces to

\[
\text{Re} \left\{ \sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} \hat{p}_e \cdot \nabla \cdot (\partial \hat{p}_e) \, d\Gamma - \sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} Z_{E} \nabla \cdot (\partial \tilde{p}_e) \, d\Gamma \right\} = 0 . \quad (6)
\]

The Stokes formula applied to \( \Omega_e \) yields

\[
\text{Re} \left\{ \sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} \left( \nabla \cdot (\partial \tilde{p}_e) - \nabla \cdot (\partial \tilde{p}_e - k^2 \tilde{p}_e - \tilde{p}_e) \right) \, d\Omega_e - \sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} Z_{E} \nabla \cdot (\partial \tilde{p}_e) \, d\Gamma \right\} = 0 . \quad (7)
\]

Therefore, the real part of (7) yields

\[
- \sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} 2
\]

\[
\sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} \left( Z_{E} \nabla \cdot (\partial \tilde{p}_e) \cdot \nabla \cdot (\partial \tilde{p}_e) \right) \, d\Gamma = 0 . \quad (8)
\]

This equation shows that \( \tilde{p}_e \) equals zero provided the absorption coefficients are nonzero, which proves the uniqueness of the solution of (4) and, consequently, the equivalence between (1) and (4).

One can see that Problem (3) and (4) is a weak formulations of the reference problem (2) which uses power quantities and satisfies the boundary conditions of (2) in a weak sense. This is a mixed asymmetrical formulation which differs from the classical symmetrical Galerkin formulation and which can be applied to ill-posed problems: it is possible to prescribe different types of boundary conditions on the same boundary. For example, in order to prescribe both the pressure \( p_{ad} \) and the velocity \( v_{ad} \) over the boundary \( \partial p_{ad}, \Omega_e \), one must use in (4) the following simple contribution:

\[
\text{Re} \left\{ \sum_{e=1}^{N_e} \int_{r_{e} \cup \partial \Omega} \left( \nabla \cdot (\partial \tilde{p}_e) \right) \cdot \left( \nabla \cdot (\partial \tilde{p}_e) \right) \, d\Gamma \right\} . \quad (9)
\]

2.3. Approximation spaces

In order to find approximate solutions of the VTCR formulations (4), one seeks \( p_E^k \) in the form \( p_E^k = \tilde{p}_E^k + p_P^k \) in each cavity \( \Omega_e \), where \( \tilde{p}_E^k \), as discussed before, denotes a known particular solution and \( p_P^k \) is taken in the discretized admissible space \( S_{ad,0} \).

Indeed, the approximations in the subcavities can be independent of one another. Such flexibility may not exist with methods which use Lagrange multipliers to assemble shape functions into a global approximation. It is important to note that (4) leaves great flexibility in the selection of the space \( S_{ad} \) used in the approximation.Indeed, the appropriate choice of the lagrange multiplier representation is a question which must also be addressed for every new set of interior shape functions. In the case of the VTCR, different families of \( S_{ad,0} \) can be used to obtain an efficient approximation of the solution.
All the previous works used for $S_{ad}$ the space of the waveband functions:

$$
S_{ad} = \text{span} \left\{ \sum_{n=0}^{N_E-1} e^{ik_n x} \cos \phi_n \sin \theta_n d\theta, n = 0, \ldots, N_E - 1 \right\}.
$$

This space is the set of the linear waveband combinations obtained as the superposition of all the plane waves traveling in directions $(\cos \theta_n, \sin \theta_n)$ and $(\cos \theta_n, \sin \theta_n)$. These functions are good candidates for the representation of a diffuse acoustic field in a finite element context. Some examples of real parts of such waveband functions are shown in Fig. 2.

In this paper, we use a different space: the space of the functions based on Fourier series:

$$
\mathcal{A}^{Eh} = \text{span} \left\{ \int_{-\pi}^{\pi} e^{ik x \cos \alpha_y \sin \theta} d\theta, n = -N_E, \ldots, N_E \right\}.
$$

The space of the waveband functions approximates the amplitudes $A_{eh}^E$ as piecewise polynomials, whereas the space of the Fourier-based functions approximates these amplitudes as truncated Fourier series. Some examples of other solutions of such Fourier-based functions are shown in Fig. 3.

One can observe that both spaces $S_{ad}^{Eh}$ and space $\mathcal{A}^{Eh}$ are included in $S_{ad}^{S}$. (The plane wave $e^{ik x \cos \alpha_y \sin \theta} d\theta$ satisfies the Helmholtz equation (3).) These two spaces are of the general form $p_{eh}^E(x) = \sum_{n=-N_E}^{N_E} A_{eh}^E(\theta) e^{ik x \cos \alpha_y \sin \theta} d\theta$, where $A_{eh}^E(\theta)$ is either a piecewise constant function (for the waveband functions) or a Fourier series term (for the Fourier series functions). Furthermore, the superscript $h$ which appears in $A_{eh}^E(\theta)$ alone shows that the VTC keeps only the amplitudes of the waves, not their spatial shapes. In that sense, the VTC belongs to the category of the multiscale numerical approaches because it discretizes only the slowly oscillating quantities and not the rapidly oscillating quantity $e^{ik x \cos \alpha_y \sin \theta} d\theta$.

Finally, one can note that all the propagating waves are taken into account since in our approximation they are represented by an integral over the space variable (and not by a discrete sum which would select only a few propagation directions). This is important because the direction of the propagation of the waves is known a priori.

Studies of the space of the waveband functions (10) have already been presented in [33,34]. These works showed the great interest of these functions for the solution of medium-frequency problems. The objective of the present paper is to analyze the efficiency of the Fourier-based approach (11) compared to the waveband-based approach and to the classical FEM.

### 2.4. Definition of an error indicator

Since the reference problem (1) is discretized, it is necessary to evaluate the accuracy of the approximate solution. The local error can be measured by:

$$
\epsilon_E = \left(\frac{E_{exact}(\Omega_E)}{\sum_{E} E_{exact}(\Omega_E)}\right)^{1/2}.
$$

where $E_{exact}(\Omega_E) = \frac{1}{m} \sum_{E} E_{exact}(\Omega_E)$ is the dissipated energy, $mes(\Omega)$ and $mes(\Omega)$ denote the measures of $\Omega$ and $\Omega_E$ respectively, and $p_{eh}^E$ corresponds to the exact solution of the problem in $\Omega_E$. (Due to energy conservation, the dissipated energy is equal to the injected energy.) One can see that this error measures the relative difference between the approximate solution and the exact solution in terms of dissipated energy. The dissipated energy is interesting in the medium-frequency range because it is a relevant quantity at such frequencies; a more local quantity, such as the pressure at a point, would be too sensitive to small variations in the problem's data.

Since the exact solution $p_{eh}^E(x)$ satisfies all the boundary conditions (including the continuity between adjacent subcavities) exactly for each subcavity, (12) constitutes a measure of the error in all the boundary conditions, including the possible discontinuities with the adjacent pressure field. Therefore, it is possible to derive a global indicator which includes all the possible discontinuities of the solution between subcavities by maximizing all the global indicators; if this indicator is zero, all the boundary conditions (including the continuity between subdomains) are satisfied.

$$
\epsilon = \max_{E} \epsilon_E
$$

The problem with the actual error (12) is that in general the exact solution is not known. Therefore, one needs to define an error estimator. This is not an easy task because some subcavities $\Omega_E$ may not touch the boundary $\partial \Omega$. For such subcavities, it is impossible to check the satisfaction of the boundary conditions, which are the only equations of the problem that are not satisfied automatically. (Remember that the shape functions satisfy the Helmholtz equation.) The only way to evaluate the accuracy of the approximate solution in each subcavity is to verify the continuity in terms of pressure and velocity with all the other subcavities in

![Fig. 3. Examples of real parts of complex pressure fields associated with Fourier-based functions (11). $\Omega_k = [0 \text{ m} 0.5 \text{ m}] \times [0 \text{ m} 0.5 \text{ m}]$, $k = 120 \text{ m}^{-1}$, $N_e = 30$, $n = 0$ (left), $n = 12$ (center) and $n = 23$ (right).](image-url)
the vicinity of \( \Omega_e \). But this verification is hampered by the fact that the solutions in the surrounding subcavities are only approximate solutions.

Based on the previous remarks, our proposed local error estimator of (12) is

\[
E_h^b = \frac{E_{d,\Omega_e}(p^b - p^{E_h^b})/\text{mes}(\Omega_e)}{\sum E_{d,\Omega_e}(p^b - p^{E_h^b})/\text{mes}(\Omega)}.
\]

(14)

where \( p^b \) corresponds to the solution of the problem in \( \Omega_e \) when the pressure and velocity at the boundaries of \( \Omega_e \) are prescribed such that they correspond to the pressure and velocity in all the \( \Omega_e \) adjacent to \( \Omega_e \). Let us recall that both the pressure and the velocity can be prescribed on the same boundary thanks to (9).

Thus, the indicator is a measure of the compatibility of the two pressure fields on both sides of the interior edge. Moreover, the solution set \( p_h^{\text{ref}} \) is calculated using a more refined discretization of \( S_{\text{ref}} \) (in practice, 10 shape functions more for \( p^b \) than for \( p^h \)). Then, if there is only one subdomain (i.e., \( n_h = 1 \)), error indicator (14) is a measure of the relative error with respect to a more refined solution used as the reference.

In order to verify the effectiveness of the error estimator (14), let us consider the following example: \( \Omega \) is a \([2.3 \text{ m} \times 1 \text{ m}]\) rectangular cavity with a curved upper-right corner (diameter 1.6 m). This cavity is filled with air \((\rho_0 = 1.25 \text{ kg m}^{-3}, c_o = 340 \text{ m s}^{-1})\) and \( \eta = 10^{-5} \) and the boundary conditions are such that the exact solution is

\[
p_e(x) = \int_{1-\pi}^{1-\pi} A^{(n)}(\theta) \cdot e^{ik(x \cdot \cdot - \theta)} d\theta,
\]

with \( A^{(n)}(\theta) = (3+i)(\theta - \pi)^1(\theta - \pi)^2 + 2 \cos \theta \cdot \sin \theta \cdot \cos \theta \cdot x_c = \{0.62, 0.42\} \), and \( k \) the wave vector of the problem (1). Such a solution is interesting because it cannot be expanded as a Fourier series easily. The boundary conditions and the solution are shown in Fig. 4. In order to illustrate the general case, all the types of boundary conditions are represented (prescribed pressure, prescribed velocity and a combination of the two, i.e. a Robin condition). Fig. 4 also shows the decomposition of \( \Omega \) into 9 subcavities of different sizes and shapes such that some touch \( \partial \Omega \) and others do not. Two circular frequencies \( k = 30 \text{ m}^{-1} \) and \( k = 90 \text{ m}^{-1} \) were chosen; thus, the characteristic length of \( \partial \Omega \) was about 10 and 30 wavelengths, respectively.

The actual local error (12), the local error estimator (14) and the relative error \( H^1(\Omega_e) \) in each of the subdomains are compared in Fig. 5. (Here, the exact solution is known).

In Fig. 5, the indicator based on the dissipated energy that the accuracy is good at relatively high error levels (i.e. at levels greater than \( 10^{-5} \)). Then, while the H1 error stops decreasing, the indicator begins oscillating as the number of degrees of freedom increases. These oscillations occur at levels from about \( 10^{-5} \) to \( 10^{-6} \). These two types of behavior below \( 10^{-5} \) can be explained by the fact that the two errors were calculated using two different reference solutions: the exact solution of the problem for the H1 error, and \( p_h^{E_h} \) for the error indicator (see Eq. (14)). The solution \( p_h^{E_h} \) was calculated at each point of the convergence curve using a more refined discretization. Thus, the reference solution changed continuously for the error indicator, but not for the H1 error. This may explain the small oscillations which can be observed for the error indicator.

The reliability of this error measure justifies its use in deriving the error estimator (14), which is confirmed in Fig. 5.

2.5. Remarks and implementation aspects

To conclude this section, some remarks must be made:

1. Seeking an approximate solution \( p^b \in S_{\text{ref}}^h \) to the variational formulation (4) requires the resolution of the finite-dimension matrix system

\[
\begin{align*}
K_{11} & \cdots K_{1n_b} \cdots & a_1 & = & b_1 \\
\vdots & \cdots & \vdots & \cdots & \vdots \\
K_{n_b1} & \cdots K_{n_bn_b} \cdots & a_{n_b} & = & b_{n_b}
\end{align*}
\]

(15)

where \( K_i \) and \( b \) are respectively the matrices of the bilinear forms and the vectors of the linear form of the variational formulation (4) calculated for all the test functions in all the subcavities \( \Omega_e \), \( a_i \) is the vector of the amplitudes of the shape functions used in \( \Omega_e \).

2. Since the shape functions used in (10) or in (11) are defined over the entire acoustic subcavity \( \Omega_e \), the matrices of the model are fully populated. However, in the case of more than two cavities, if Cavity \( i \) is not connected to Cavity \( j \) the corresponding off-diagonal submatrix \( K_{ij} \) is zero.

3. Concerning variational formulation (4) without discretization, one can prove that in the absence of dissipation and if the circular frequency is not an eigenvalue one obtains the classical solution. If the circular frequency is an eigenvalue, the VTCR has a solution if, and only if, the work associated with the eigenmodes equals zero. The use of formulation (3) with discretization leads to an approximate eigenvalue problem for which similar results are obtained numerically. These particular cases will be studied in an upcoming paper.

![Fig. 4. Top: description of the example. Bottom: the real part of the exact pressure field for \( k = 30 \text{ m}^{-1} \) (left) and \( k = 90 \text{ m}^{-1} \) (right).](image)
4. A preconditioner can be associated with the VTCR. In the following examples, this preconditioner is defined as the diagonal matrix of the inverse of the square root of the energy of the shape functions. Consequently, all the shape functions are approximately equivalent in terms of energy and, thus, the operator is calculated using what can be viewed as energy-normalized shape functions. This makes their contributions to the solution easier to compare. Then, thanks to Moore Penrose’s pseudo-inverse, the scaled matrix is inverted using a direct solver.

5. Since the shape functions used in (10) or in (11) depend implicitly on the frequency, so do the matrices of the model: therefore, these cannot be partitioned into frequency-independent submatrices. However, strategies for calculating the solution over a range of frequencies without calculating the response at each frequency have been developed (see [30]).

6. Once the discretization has been chosen, the calculation of the matrix coefficients involves the integration of highly oscillatory functions. Special care must be taken in carrying out these integrations in order to ensure proper convergence. Referring to Eq. (11) and (4), a typical term to be calculated is:

\[ \int_{\omega F} \int_{\theta \omega \pi} \int_{\phi \omega \pi} e^{im\theta(k\sin \theta)} e^{-im\phi} e^{-ik\theta} e^{-ik\phi} d\theta d\phi d\omega. \]

If \((k \cos \theta, k \sin \theta)\) and \((r \cos \phi, r \sin \phi)\) denote the Cartesian coordinates of \(k(\omega)\) and \(x\), respectively, a typical term to be calculated becomes

\[ \int_{\omega F} \int_{r \omega \infty} \int_{\phi \omega \pi} e^{im(r \cos \phi + \sin \phi)} e^{-im\phi} e^{-ik\theta} e^{-ik\phi} d\phi d\theta dr, \]

which, using the Anger–Jacobi expansion \(e^{im\cos \phi} = \sum_{n=-\infty}^{\infty} J_n(kr)(-i)^n e^{-in\phi} \), where \(J_n(\cdot)\) is the \(n\)th Bessel function and taking into account the orthogonality of the \(e^{im\phi}\) terms, leads to \(4\pi^2 \int_{\omega F} \int_{r \omega \infty} e^{im\phi} J_n(kr)(-i)^n e^{-in\phi} d\phi dr\). Then, quadratures need to be performed only on the boundary \(F\), which reduces the computational cost of the method dramatically. Numerical integration schemes such as the classical Simpson's integration with frequency-dependent numbers of quadrature points can be used (see [35]). Graphics Processing Units (GPU) or multiple parallel servers can also be used to accelerate many of the computations without resorting to low-level programming.

7. A numerical advantage of Fourier-based functions comes from the hierarchical nature of the approximation space, which leads to the inclusion \(F_{ad}^{E_{ad}} \subset F_{ad}^{E_{ad}-1}\) (i.e., the approximation space (11), of dimension \(2N_E + 1\), is included in the approximation space of dimension \(2N_E + 3\)). That is not the case with waveband functions (10). A direct consequence is that matrix \(K_{ad}\), of dimension \((2N_E + 1) \times (2N_E + 1)\), can be used to calculate matrix \(K_{E_F}\), of dimension \((2N_E + 3) \times (2N_E + 3)\). Thus, the size of the Fourier approximation space \(S_{ad}\) can be increased to improve convergence at no significant additional cost.

8. The presence of a non-zero right-hand term \(f(x)\) in the Helmholtz Eq. (1) can be taken into account simply by adding the associated particular solution \(P_E(x) = \int_{\omega F} f(x) G(x,y) dS(y)\), where \(G(x,y)\) is the Helmholtz equation's Green function, to approximation spaces \(S_{ad}\) (see [27] for further details and some numerical examples).

### 3. Advantages of the Fourier-based VTCR

#### 3.1. Improvement in the robustness of the method

The Fourier VTCR shape functions \(\int_{\omega F} e^{im\phi} \cos(k \cos \phi + \sin \phi) d\phi\) behave differently than the waveband VTCR shape functions \(\int_{\omega F} e^{im\phi} \cos(k \cos \phi + \sin \phi) d\phi\). All the waveband VTCR shape functions in the approximation space \(W_{E_F}\) have similar energy levels: they have almost the same form and differ only in the directions of propagation of the waves. Conversely, the energy of the Fourier VTCR shape functions decreases with \(n\). This can be observed in Fig. 6 and is further illustrated in Fig. 3, where one can clearly see that the
maximum amplitude of the shape functions decreases with \( n \). Then, as one approaches a numerical accuracy \( e \), the Fourier VTCR increases the size of \( \mathcal{F}_E \); had with new functions whose energy contribution is less than \( e \). Since these additional functions cannot be calculated properly, the size of \( \mathcal{F}_E \) must be limited to the shape functions whose energy contribution is greater than the numerical accuracy. Moreover, in order to avoid numerical difficulties in the resolution of the system, one uses a preconditioner such as the diagonal matrix of the inverse of square root of the energies of the shape functions, which amounts to computing the operators with energy-normalized shape functions.

Furthermore, Fig. 7 shows the actual error of the Fourier VTCR in each subcavity as a function of the energy of the last shape functions at \( k = 30 \text{ m}^{-1} \). In this figure, one can clearly see that an acceptable level of accuracy is achieved when the energy of the last shape function is less than \( 10^{-3} \) for each subdomain considered. This result can be used in order to estimate \textit{a priori} the number of DOFs which is necessary for a reliable calculation and a good solution.

Finally, Fig. 8 shows, on the same diagram, the convergence curves of the actual error (12) (which, in this particular example, is known) and the conditioning numbers of the VTCR matrix \( \mathbf{K} \) at \( k = 30 \text{ m}^{-1} \), both for the waveband and for the Fourier VTCR with the preconditioner. These results present the same behavior as that already observed in many previous works (see [33,37,34]): when the number of shape functions becomes too large, the conditioning number of the numerical matrix deteriorates and convergence is interrupted. This is due to the limited accuracy of the computer, which is unable to distinguish two very similar shape functions. This behavior
was observed in other numerical strategies using oscillating functions in the approximation spaces (see, e.g. [36,18]). Some means of reducing this phenomenon were proposed in [37], but they all involve expensive matrix manipulations. Nevertheless, one can see in the previous figures that the conditioning number of the numerical matrix is better in the case of the Fourier series, and the convergence stop due to numerical precision happens for a larger number of DOFs than the waveband approach. Indeed, for the same accuracy, the conditioning number of the Fourier VTCR is less than that of the waveband VTCR. One of the advantages of the Fourier approximation is that no regularization is required.

3.2. Some remarks on the efficiency of the Fourier approach

Now, let us consider the application of the Fourier VTCR to the more complex example of Fig. 9. The cavity is filled with air ($\rho_0 = 1.25 \text{ kg m}^{-3}$, $c_0 = 340 \text{ m s}^{-1}$ and $\eta = 10^{-5}$) and Robin, Dirichlet and Neumann boundary conditions are prescribed along different edges. Two different meshes were used in order to compare the results of the VTCR with a uniformly sized mesh (mesh A) and with an unstructured mesh (mesh B). The two frequencies chosen were $1800 \text{ Hz}$ and $4800 \text{ Hz}$, which correspond respectively to 7 and 18 wavelengths in the cavity. For each frequency, the Fourier VTCR results were compared to two FEM solutions obtained with standard $Q_2$ and $Q_3$ elements and a regular mesh of size $h$. The characteristic length $h$ was chosen so that the classical rule of 10 elements per wavelength would be respected. The VTCR and FEM solutions are shown in Fig. 10.

Fig. 11 shows the convergence curves of the Fourier VTCR and the $Q_2$ and $Q_3$ FEM with respect to the number of DOFs for the two frequencies considered. In this case, mesh A alone is used (see Fig. (9)). In the absence of an exact analytical solution, the convergence of the VTCR and the FEM convergence are calculated using (12) and assuming that an overkill FEM solution of the problem represents the exact solution $p^{\text{ex}}$. According to that figure, all three methods (Fourier VTCR and $Q_2$ and $Q_3$ FEM) converge, but the VTCR method converges faster by more than an order of magnitude. Indeed, for the VTCR, the convergence rate appears to be exponential. This good convergence rate was already observed in the case of the waveband VTCR (see [33]). The comparison of the VTCR and FEM curves clearly illustrates the greater efficiency of the numerical methods which use plane waves in the approximate resolution of medium-frequency problems.

The computation costs of the Fourier VTCR and the FEM cannot be easily compared because the solvers were different and the authors used a commercial FEM program which may not be fully optimized. However, a simple comparison of the total times for integrating, assembling and solving the equations on a MacBook computer with a 2.4 GHz Intel Core 2 duo processor and 4 GB RAM shows that despite the need to integrate rapidly oscillating functions the VTCR program performs very well compared to the commercial FEM code. Indeed, in all our computations, we observed that the VTCR runs as fast as the FEM, or even faster, for the higher-frequency test cases. An explanation can be found in the remark of Section 2.5 concerning the semi-analytical method used in the VTCR.

An important numerical consideration is the choice of the discretization of the spatial decomposition into subdomains and of the size of $S^m$. An empirical rule for the determination of the optimum number of waveband shape functions to be used in each

![Fig. 9. Definition of the L-shaped cavity. Top: the applied boundary conditions. Bottom: the two meshes, denoted A and B.](image_url)

![Fig. 10. The real parts of the VTCR, FEM $Q_2$ and FEM $Q_3$ solutions of the problem of 9.](image_url)
subcavity, based on a geometrical criterion, was investigated in [37]. According to that rule, the size of $W_{ad}$ must be proportional to the ratio of the perimeter of the circumscribed circle to the wavelength. A ratio equal to 4 guarantees good accuracy. Here, we present an alternative criterion based on the remarks of the last paragraph of Section 3.1 and illustrated in Fig. 7. We showed that the magnitude of the error is related to the number of Fourier VTCR DOFs being used and that it is unnecessary to enrich the space $\mathcal{F}_h$ with functions that are negligible in terms of energy. Then, looking at the energies of the Fourier VTCR shape functions in each subcavity, it is easy to select the right number of functions to achieve the same magnitude of the error in all subcavities. Based on the considerations of Section 3.1, we selected an energy ratio of $10^{-2}$ between the last and the first shape functions retained. Fig. 12 shows a comparison of the waveband and Fourier VTCR applied to mesh A and mesh B. One can see that convergence is achieved with fewer DOFs in mesh A than in mesh B, which is consistent with previous works on waveband VTCR (see [33,37]) and seems to indicate that the VTCR is more efficient in the case of a small number of large subdomains. Furthermore, it seems that the energy criterion, besides being physically more meaningful, is more accurate than the geometrical criterion in predicting the optimum number of DOFs.

3.3. Comparison with benchmark examples

In order to assess the performance of the VTCR, we propose to use the same two benchmark examples as in [38]. These examples were chosen in [38] for assessing the respective performances of the Galerkin, Galerkin least squares and residual free bubbles methods.

First, let us consider a square domain $\Omega = (0, L) \times (0, L)$ with homogeneous Dirichlet boundary conditions. The force load is a Dirac function located at $(x_0, y_0) = (0.1875, 0.1875)$ and the nondimensional wave numbers are $kL = 8, 24, 40, 56, 72$ and 88. The analytical solution of this problem is given in [39]. For this example, the VTCR used a single cavity and, as pointed out in Remark 7 of Section 2.5, required the particular solution of the non-homogeneous Helmholtz equation (in this case, the pressure solution of an infinite acoustic medium subjected to the force load) to be added to the approximation space. An absorption $\eta = 10^{-3}$ was used to prevent numerical difficulties. Fig. 13 shows the VTCR solution along line $y = 0.190625$ with $kL = 8$ using 23 shape functions. Fig. 14 shows the convergence curves for all the nondimensional wave numbers considered. One should note that the VTCR was able to describe the solution both near the spike caused by the force load and away from that region quite well. The convergence curves

Fig. 11. Convergence curves for the problem defined in Fig. 9. Mesh A of Fig. 9 was used for the VTCR calculations.

Fig. 12. Convergence curves of the error estimator (14) for the problem described in Fig. 9 at 1800 Hz (left) and 4800 Hz (right). Two meshes (A and B) were used and three cases were considered: Fourier basis functions using the geometrical criterion, Fourier basis functions using the energy criterion, and waveband basis functions using the geometrical criterion.
Fig. 13. Validation using the example of a square domain with homogeneous boundary conditions, presented in [38]: the VTCR solution (continuous line) and the reference solution (dotted line) along the line $y = 0.190625$ with $kL = 8$.

Fig. 14. Validation using the example of a square domain with homogeneous boundary conditions, presented in [38]: convergence curves for nondimensional wave numbers $kL = 8, 24, 40, 56, 72$ and 88 showing the actual error (12).

Fig. 15. Validation using an L-shaped domain with various boundary conditions, presented in [38]: definition of the problem (left) and the real part of the VTCR solution for $kL = 8$ with 54 DoFs (right).

of Fig. 14 also show that for this non-homogenous Helmholtz problem the VTCR behaved quite well in different frequency ranges. Indeed, in all the cases considered, the convergence curves decreased sharply to zero.

Now, let us consider an L-shaped domain with homogeneous Dirichlet boundary conditions, except for a load applied to half of one of the faces of the L, as shown in the left part of Fig. 15. The nondimensional wave number is $k \cdot L = 8$, where $L$ is the width of the L-shaped domain. The meshes used in [38] were uniform in the three squares making up the L-shape domain. For the VTCR, the domain was divided into two symmetrical parallelepipeds, and an absorption coefficient $\eta = 10^{-5}$ was used. The energy criterion of Section 3.1 led to the choice of a total of 54 degrees of freedom. The VTCR solution is shown in the right part of Fig. 15. The
shows that Fig. 17

Comparison of the Fourier VTCR and the FEM in terms of the number of DOFs, the number of nonzero matrix coefficients and the matrix densities for the example of Table 1

two-dimensional acoustic cavity of a car, filled with air

3.4. Extension of the applicability of the VTCR to higher frequencies

Finally, we present the application of the Fourier VTCR to a two-dimensional acoustic cavity of a car, filled with air ($\rho = 1.25$ kg m$^{-3}$, $c = 340$ m s$^{-1}$ and $\eta = 10^{-5}$) and subjected to the Robin boundary conditions shown in Fig. 16. Three circular frequencies were considered: $\omega = 2\pi \times 2500$ rad s$^{-1}$, $\omega = 2\pi \times 4000$ rad s$^{-1}$ and $\omega = 2\pi \times 8000$ rad s$^{-1}$. The loading was applied along the leftmost edge. The Fourier VTCR was used inside each cavity with the optimum number of shape functions (see the remarks in the last paragraph of Section 3.1). Fig. 17 shows the calculated solutions. One can see that the solution at $\omega = 2\pi \times 8000$ rad s$^{-1}$ (the high-frequency case) contains several dozen wavelengths inside the entire acoustic cavity.

Table 1 shows a comparison of the Fourier VTCR and the FEM in terms of the number of DOFs, the number of nonzero matrix coefficients and the matrix density (i.e. the ratio between the number of nonzero terms and the size of the matrix) of the matrix. For the purpose of this comparison, the optimum number of shape functions inside each acoustic cavity (which was known a priori thanks to remarks in Section 3.1) was used for the VTCR, and 10 DOFs per wavelength were used for the FEM. (Fewer DOFs would have led to a poor solution.) From that table, one can observe the interest of the Fourier VTCR, especially at $\omega = 2\pi \times 8000$ rad s$^{-1}$. Indeed, at that frequency, the Fourier VTCR required 140 times fewer DOFs than the FEM. The number of nonzero coefficients of the VTCR matrices was about 275 times less than that of the FEM matrices. Of course, the FEM matrices density are very small as the number of DOFs is very large. Based on that observation, one can conclude that the Fourier VTCR is an efficient numerical strategy for 2D bounded acoustic problems in the mid- and high-frequency domains.

For this comparison, a relatively small $h$-refinement approach for the FEM was used, which is the most natural and commonly used practice. Additional comparisons using a $p$-refinement approach for the FEM will be made in forthcoming works.

4. Conclusion

This paper describes a new version of the VTCR for the steady-state dynamic analysis of 2D acoustic problems which uses new shape functions based on the Fourier series expansion of the amplitudes of the plane waves traveling inside the acoustic cavity. This technique is illustrated by a number of examples of varying complexity at different frequencies.

Comparisons with standard FEM and with the waveband VTCR show some interesting properties of this new version. First, the strong convergence rate which is a characteristic of the VTCR and all wave approaches is preserved with the Fourier approximation. Second, the new method is more robust and more efficient because the right number of waves to be used in each subcavity (according to an energy criterion) can be known a priori, which improves the
numerical stability. Third, the computational capabilities of the method enable it to be used at higher frequencies. Some of the problems which will have to be addressed in the future are the extension of the Fourier VTCR to 3D acoustic problems and to vibroacoustic problems.

Acknowledgements

The authors gratefully acknowledge the contribution of the ITN Marie Curie project GA-214909 “MID-FREQUENCY - CAE Methodologies for Mid-Frequency Analysis in Vibration and Acoustics”.

References