SIAM J. NUMER. ANAL.
Vol. 58, No. 3, pp. 1845–1866

FOURIER ANALYSES OF HIGH-ORDER CONTINUOUS AND DISCONTINUOUS GALERKIN METHODS*

DANIEL Y. LE ROUX†, CHRISTOPHER ELDRED‡, AND MARK A. TAYLOR§

Abstract. We present a Fourier analysis of wave propagation problems subject to a class of continuous and discontinuous discretizations using high-degree Lagrange polynomials. This allows us to obtain explicit analytical formulas for the dispersion relation and group velocity and, for the first time to our knowledge, characterize analytically the emergence of gaps in the dispersion relation at specific wavenumbers, when they exist, and compute their specific locations. Wave packets with energy at these wavenumbers will fail to propagate correctly, leading to significant numerical dispersion. We also show that the Fourier analysis generates mathematical artifacts, and we explain how to remove them through a branch selection procedure conducted by analysis of eigenvectors and associated reconstructed solutions. The higher frequency eigenmodes, named erratic in this study, are also investigated analytically and numerically.

Key words. finite element method, dispersion analysis, numerical dispersion, discontinuous Galerkin method, computational modes, erratic numerical modes

AMS subject classifications. 35L99, 74S05, 65M99, 65M12, 65T99, 74J15

DOI. 10.1137/19M1289595

1. Introduction.

In the continuum, several well-known systems, e.g., the inviscid linearized shallow-water, Maxwell, or sound wave equations, could be rewritten in the form of a set of two decoupled transport equations, yielding the dispersion relation (relation between phase speed and wavenumber) given by a constant phase solution. There is a long history of evaluating numerical discretizations by their ability to capture this behavior. For finite difference approximations, see [17, Chapter 5.3]. For high-order finite element methods, the initial work was done numerically to compute eigenvalues and eigenmodes of the discrete system [6, 7, 8, 12, 13]. Based on these results, many aspects of the discrete solutions are well understood, such as the discrete dispersion relation, group velocity, and convergence rates of these properties.

These numerical results were first confirmed analytically in [1, 2, 9], using a Bloch wave approach. The Bloch modes naturally split into two families of eigenmodes: one closely matching the low frequency solutions (up to $4h$) and the other, associated

*Received by the editors September 24, 2019; accepted for publication (in revised form) April 16, 2020; published electronically June 17, 2020.
https://doi.org/10.1137/19M1289595
Funding: The work of the first and third authors was supported by the Isaac Newton Institute for Mathematical Sciences, Cambridge, UK, where this work was initiated. The work of the second author was supported by the French National Research Agency through grant ANR-14-CE23-0010 (HEAT). Sandia National Laboratories is a multimission laboratory owned and operated by National Technology & Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent views of the U.S. Department of Energy or the United States Government.

†Université de Lyon, CNRS, Université Lyon 1, Institut Camille Jordan, 69622 Villeurbanne Cedex, France (dleroux@math.univ-lyon1.fr).

‡Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France. Current address: Center for Computing Research, Sandia National Laboratories, Albuquerque, NM 87185-1320 (chris.eldred@gmail.com).

§Center for Computing Research, Sandia National Laboratories, Albuquerque, NM 87185-1320 (mataylo@sandia.gov).

with high frequencies ranging from $4h$ up to $2h$, behaving erratically, h being the meshlength parameter. The Bloch wave approach is a powerful analytic tool, but there are reasons to develop a Fourier mode interpretation. These include the ability to derive the dispersion relation and group velocity, identify the potentially damaging emergence of discontinuities or gaps in the dispersion relation [3, 4, 12, 18], and investigate the erratic higher frequency modes, which is the goal of the present work.

The transport equation is discretized for the continuous Galerkin (CG) and discontinuous Galerkin (DG) methods using high-degree Lagrange interpolating polynomials. Contrary to previous Fourier approaches, explicit analytical formulas are provided for the dispersion relation. Consequently, the group velocity is derived and, for the first time to our knowledge, the presence of eventual gaps at specific wavenumbers is characterized analytically, and their specific locations are computed. Wave packets with energy at these wavenumbers will fail to propagate correctly, leading to significant numerical dispersion. We also show that the Fourier analysis gives rise to mathematical artifacts, and we explain how to remove them through a careful branch selection procedure driven by analysis of eigenvectors and associated reconstructed solutions. This analytical "cleaning" procedure results in a dispersion relation that is a single-valued function of wavenumber. Finally, the presence of higher frequency eigenmodes is investigated analytically and numerically. These are named erratic in this study, and they have spuriously large phase velocity errors.

The paper is organized as follows. The model problem is presented in section 2, and the CG and DG discretizations are performed in sections 3 to 4, respectively. Some concluding remarks complete the study.

2. Model problem. For an enclosed domain, let $\Omega = (0, L)$, where $0 \le x \le L$ and c is a constant parameter, and consider the transport equation

$$\frac{\partial u(x,t)}{\partial t} + c\frac{\partial u(x,t)}{\partial x} = 0 \quad \text{in } \Omega \times (0,T). \tag{2.1}$$

Boundary conditions and initial data complete the specification of (2.1). Assuming time is continuous, we seek periodic solutions of the form $u(x,t) = u(x)\, e^{-i\omega t}$, where ω is the angular frequency, and this leads to

$$i\omega u - c\frac{\partial u}{\partial x} = 0. \tag{2.2}$$

If we examine the free mode of (2.2) by perturbing about the basic state $u = 0$ and substituting a periodic solution of the form $u(x) = \widehat{u}\, e^{ikx}$ into (2.2), where $\widehat{u}$ is the Fourier amplitude and k is the wavenumber in the x-direction, we obtain

$$\omega = c\,k := \omega^{AN} \quad \text{and} \quad \frac{\partial \omega}{\partial k} = c := \frac{\partial \omega^{AN}}{\partial k} \tag{2.3}$$

for the phase speed and group velocity, respectively, and there is no dispersion.

In sections 3 and 4, (2.2) is spatially discretized using the CG and DG methods, respectively.

3. The continuous Galerkin discretization.

3.1. Continuous Galerkin formulations. We first introduce the weak formulation and then describe the CG spaces that are employed in this section. Finally, the CG discretization is presented.

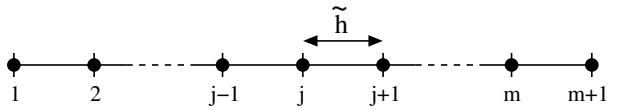

FIG. 1. *The position of nodes* $j = 1, 2, 3, \ldots, m+1$.

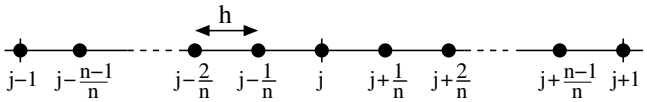

FIG. 2. *Indices of the local degrees of freedom on element* e_j *of* ε_h, $j = 1, 2, \ldots, m$.

The space $H^1(\Omega)$ is the Sobolev space of functions in the square-integrable space $L^2(\Omega)$, whose first derivatives belong to $L^2(\Omega)$. Let u be in a subspace V of $H^1(\Omega)$. Multiplying (2.2) by a function v belonging to V and integrating over Ω yields

$$i\omega \int_\Omega uv\,dx - c\int_\Omega \frac{\partial u}{\partial x}\,v\,dx = 0 \quad \forall v \in V. \tag{3.1}$$

Let ε_h be a partition of Ω, and let $\widetilde{h}$ be the size of an element; namely, ε_h is a finite collection of m elements e_j, $j = 1, 2, \ldots, m$, of the real line, such that $\bar{\Omega} = \bigcup_{e_j \in \varepsilon_h} \bar{e}_j$. Consider a mesh of m intervals of length $\widetilde{h}$ on Ω with elements $e_j = (x_j, x_{j+1})$ and nodes x_j for $j = 1, 2, \ldots, m$, as shown in Figure 1. For the sake of performing the subsequent Fourier analyses in subsection 3.2, periodicity of the solution is imposed at end nodes $j = 1$ and $j = m+1$, and we have $x_1 = x_{m+1}$. For $n \in \mathbb{N}$, finite element spaces V_h^n of polynomial functions are considered, continuous at the element interfaces, such that $V_h^n \subset V$, with

$$V_h^n = \{\, u_h \in H^1(\Omega);\ u_h|_e = \widetilde{u}_h \circ F_e^{-1},\ \widetilde{u}_h \in \mathcal{P}_n(\widetilde{e})\ \forall\, e \in \varepsilon_h \},$$

where F_e is the affine mapping from the master element $\widetilde{e}$ to the element e in the partition ε_h, and $\mathcal{P}_n(\widetilde{e})$ is the space of polynomial functions of degree at most n on $\widetilde{e}$. In order to match the requirements of the Fourier analysis, the mesh is uniform and Lagrange test functions are employed. The basis $\mathcal{V}_h^n$ of V_h^n is chosen such that on e_j, $\mathcal{V}_h^n$ is a basis of the $n+1$ Lagrange interpolating functions of degree n with $\mathcal{V}_h^n = \{v_s\}$ for $s \in \mathcal{I}_j^n$, $j = 1, 2, \ldots, m$, where $\mathcal{I}_j^n = \left\{j, j+\frac{1}{n}, j+\frac{2}{n}, j+\frac{3}{n}, \ldots, j+\frac{n-1}{n}, j+1\right\}$ is the set of indices of the local degrees of freedom on each element e_j of ε_h, $j = 1, 2, \ldots, m$, as shown in Figure 2 for elements e_{j-1} and e_j. Further, since the mesh is uniform, we let $x_{j+\frac{q+1}{n}} - x_{j+\frac{q}{n}} = h = \widetilde{h}/n$ for $j = 1, 2, \ldots, m$ and $q = 0, 1, 2, \ldots, n-1$. Introducing the finite element basis leads to a finite element discretization of (3.1), and at node s, we search for u_h belonging to V_h^n for the selected basis such that

$$i\omega \sum_{j=1}^{m} \int_{e_j} u_h v_s\,dx - c\sum_{j=1}^{m}\int_{e_j} \frac{\partial u_h}{\partial x}\,v_s\,dx = 0 \quad \forall v_s \in V_h^n. \tag{3.2}$$

For a typical node j, $j = 1, 2, \ldots, m$, we let

$$\widetilde{\omega} = \frac{\omega h}{c}, \quad M_{q,r} = \frac{1}{h}\int_{e_j} v_{j+\frac{q}{n}}\, v_{j+\frac{r}{n}}\,dx, \quad G_{q,r} = \int_{e_j} v_{j+\frac{q}{n}}\,\frac{\partial v_{j+\frac{r}{n}}}{\partial x}\,dx \tag{3.3}$$

for $0 \le q, r \le n$, where $\widetilde{\omega}$ is the normalized frequency and $M_{q,r}$ and $G_{q,r}$ are the elements of the local mass and gradient matrices of size $(n+1)\times(n+1)$, denoted by $\mathcal{M}$ and $\mathcal{G}$, respectively, with

$$\mathcal{M} = (M_{q,r}), \qquad \mathcal{G} = (G_{q,r}) \quad \text{for } 0 \le q, r \le n. \tag{3.4}$$

The properties of the Lagrange basis functions, in particular $\sum_{q=0}^{n} v_{j+\frac{q}{n}} = 1$, lead to

$$M_{q,r} = M_{n-q,n-r} \quad \text{and} \quad G_{q,r} = -G_{n-q,n-r} \quad \text{for } 0 \le q, r \le n, \tag{3.5}$$

$$\sum_{r=0}^{n} G_{q,r} = 0 \quad \text{for } 0 \le q \le n \quad \text{and} \quad \sum_{q=0}^{n} G_{q,r} = 0 \quad \text{for } 1 \le r \le n-1, \tag{3.6}$$

$$\sum_{q=0}^{n} G_{q,0} = -1, \quad \sum_{q=0}^{n} G_{q,n} = 1, \quad \text{and} \quad G_{0,0} = -G_{n,n} = -1/2. \tag{3.7}$$

In order to perform the subsequent Fourier analyses, we need to compute (3.2) at nodes $s = j \pm \frac{q}{n}$, $q = 1, 2, 3, \ldots, n-1$, and $s = j$, $j = 1, 2, \ldots, m$. Expanding u_h over e_{j-1} and e_j in the basis $\mathcal{V}_h^n$ for $j = 1, 2, \ldots, m$ and using (3.5) yields

$$\sum_{r=0}^{n} u_{j\pm\frac{r}{n}} \left(i\widetilde{\omega} M_{q,r} \mp G_{q,r}\right) = 0 \quad \text{at } s = j \pm \frac{q}{n}, \quad q = 1, 2, 3, \ldots, n-1, \tag{3.8}$$

$$\sum_{r=0}^{n} \left(u_{j-\frac{r}{n}} (i\widetilde{\omega} M_{0,r} + G_{0,r}) + u_{j+\frac{r}{n}} (i\widetilde{\omega} M_{0,r} - G_{0,r}) \right) = 0 \quad \text{at } s = j. \tag{3.9}$$

Due to periodicity at end nodes, with $u_0 = u_m$ and $u_{m+1} = u_1$, (3.9) yields m equations and (3.8) provides $m(n-1)$ additional equations at internal element nodes, leading to a total of mn discrete equations.

3.2. The Fourier analysis. A dispersion analysis is now performed, and periodic solutions of (3.8) and (3.9) are sought. For symmetry reasons, only selected discrete equations are considered, namely $n-1$ equations at interior nodes $j + \frac{q}{n}$ on element e_j obtained from (3.8) for $q = 1, 2, \ldots, n-1$ and one equation at a boundary element node j obtained from (3.9). The results of the present section are independent of the choice of the typical node j, $j = 1, 2, \ldots, m$. The periodic solutions read as follows:

(i) $u_{j-\frac{r}{n}} = \widehat{u}_{n-r}\, e^{ikx_{j-\frac{r}{n}}}$ at the $n-1$ interior nodes $j - \frac{r}{n}$ of e_{j-1}, $r = 1, 2, \ldots, n-1$;

(ii) $u_{j+\frac{r}{n}} = \widehat{u}_r\, e^{ikx_{j+\frac{r}{n}}}$ at the $n-1$ interior nodes $j + \frac{r}{n}$ of e_j, $r = 1, 2, \ldots, n-1$;

(iii) $u_j = \widehat{u}_n\, e^{ikx_j}$ and $u_{j\pm1} = \widehat{u}_n\, e^{ikx_{j\pm1}}$ at the boundary nodes of e_{j-1} and e_j.

For symmetry reasons, the same Fourier amplitudes $\widehat{u}_r$ are considered at nodes $j - \frac{n-r}{n}$ of e_{j-1} and at nodes $j + \frac{r}{n}$ of e_j, $r = 1, 2, \ldots, n-1$. Similarly, the amplitudes at end nodes $j-1$, j, and $j+1$ of e_{j-1} and e_j are identical and are denoted by $\widehat{u}_n$. A total of n amplitudes are hence involved, as shown in Figure 3. For $q = 1, 2, \ldots, n-1$, substituting $u_{j+\frac{r}{n}}$, u_j, and u_{j+1} in (3.8) leads to

$$\sum_{r=1}^{n-1} \widehat{u}_r\, e^{irkh} \left(i\widetilde{\omega} M_{q,r} - G_{q,r}\right) + \widehat{u}_n \left(i\widetilde{\omega} M_{q,0} - G_{q,0} + e^{inkh} \left(i\widetilde{\omega} M_{q,n} - G_{q,n}\right)\right) = 0, \tag{3.10}$$

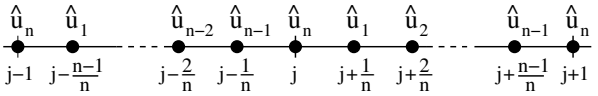

FIG. 3. *Indices of the Fourier amplitudes on element e_j of ε_h, $j = 1, 2, \ldots, m$.*

and substituting $u_{j\pm\frac{r}{n}}, u_j$, and $u_{j\pm1}$ in (3.9) using (3.5) yields

$$\sum_{r=1}^{n-1} \widehat{u}_r \Big(e^{irkh} \left(i\widetilde{\omega} M_{0,r} - G_{0,r}\right) + e^{-i(n-r)kh} \left(i\widetilde{\omega} M_{n,r} - G_{n,r}\right) \Big) + \widehat{u}_n \Big(2i\widetilde{\omega}(M_{0,0} + \cos(nkh) M_{0,n}) - 2i \sin(nkh) G_{0,n} \Big) = 0. \tag{3.11}$$

Equations (3.10)–(3.11) yield the matrix system for the Fourier amplitudes

$$\left(i\widetilde{\omega}\widehat{\mathcal{M}} - \widehat{\mathcal{G}}\right)\widehat{\mathbf{U}} = 0, \tag{3.12}$$

where $\widehat{\mathbf{U}} = (\widehat{u}_1, \widehat{u}_2, \ldots, \widehat{u}_n)$, $\widehat{\mathcal{M}}$ and $\widehat{\mathcal{G}}$ are the $n \times n$ matrices, such that $\widehat{\mathcal{M}} = \widehat{\mathcal{N}}_M$ with $\alpha = 2(M_{0,0} + \cos(nkh) M_{0,n})$ and $\widehat{\mathcal{G}} = \widehat{\mathcal{N}}_G$ with $\alpha = 2i \sin(nkh)\, G_{0,n}$, and $\widehat{\mathcal{N}}_N$ reads as

$$\widehat{\mathcal{N}}_N = \begin{pmatrix} N_{1,1} & \cdots & N_{1,n-1} & N_{1,0} + e^{inkh} N_{1,n} \\ \vdots & \ddots & \vdots & \vdots \\ N_{n-1,1} & \cdots & N_{n-1,n-1} & N_{n-1,0} + e^{inkh} N_{n-1,n} \\ N_{0,1} + e^{-inkh} N_{n,1} & \cdots & N_{0,n-1} + e^{-inkh} N_{n,n-1} & \alpha \end{pmatrix}.$$

For a nontrivial solution $\widehat{\mathbf{U}}$ to exist, $\widetilde{\omega}$ has to satisfy the characteristic equation or dispersion relation $\det(i\widetilde{\omega}\widehat{\mathcal{M}} - \widehat{\mathcal{G}}) = 0$, namely a polynomial equation of degree n.

THEOREM 3.1. *The following properties hold:*

(i) *$\widehat{\mathcal{G}}$ is skew-Hermitian and $\widehat{\mathcal{M}}$ is Hermitian.*

(ii) *$\widehat{\mathcal{M}}$ is a positive definite matrix.*

(iii) *The roots $\widetilde{\omega}$ of the characteristic equation $\det(i\widetilde{\omega}\widehat{\mathcal{M}} - \widehat{\mathcal{G}}) = 0$ are real.*

Proof. Let $\mathcal{X}^*$ be the conjugate transpose of a matrix $\mathcal{X}$. Property (i) holds due to (3.5) and the definition of α. To prove (ii), let $\mathcal{P}$ be the $(n+1) \times (n+1)$ matrix whose n first columns are the n first columns of the $(n+1) \times (n+1)$ identity matrix, and let the last column of $\mathcal{P}$ read as $(1, 0, \ldots, 0, e^{inkh})^T$. Since $\mathcal{M}$ is positive definite [5] and $\det \mathcal{P} = e^{inkh} \neq 0$, the matrix $\mathcal{P}^*\mathcal{M}\mathcal{P}$ is also positive definite. Indeed, for all $\mathbf{v} \in \mathbb{C}^{n+1}$ we have $\mathbf{v}^*\mathcal{P}^*\mathcal{M}\mathcal{P}\mathbf{v} = (\mathcal{P}\mathbf{v})^*\mathcal{M}(\mathcal{P}\mathbf{v})$, and because $\mathcal{M}$ is a real matrix we obtain $\mathbf{v}^*\mathcal{P}^*\mathcal{M}\mathcal{P}\mathbf{v} \in \mathbb{R}$. Elementary computations then show that $\widehat{\mathcal{M}}$ is nothing more than a leading principal submatrix of $\mathcal{P}^*\mathcal{M}\mathcal{P}$ obtained from $\mathcal{P}^*\mathcal{M}\mathcal{P}$ by removing the first column and the first row. Consequently, $\widehat{\mathcal{M}}$ is also a positive definite matrix.

Finally, the generalized eigenvalue problem (3.12) is rewritten as $i\widehat{\mathcal{G}}\,\widehat{\mathbf{U}} = -\widetilde{\omega}\widehat{\mathcal{M}}\,\widehat{\mathbf{U}}$, where $i\widehat{\mathcal{G}}$ and $\widehat{\mathcal{M}}$ are both Hermitian matrices due to property (i). Since $\widehat{\mathcal{M}}$ is definite positive, the roots $\widetilde{\omega}$ of the characteristic equation $\det(i\widehat{\mathcal{G}} + \widetilde{\omega}\widehat{\mathcal{M}}) = 0$ are real and two eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ with distinct eigenvalues are $\widehat{\mathcal{M}}$-orthogonal with $\mathbf{v}_1^*\widehat{\mathcal{M}}\mathbf{v}_2 = 0$. Further, there exists a basis of generalized eigenvectors. ∎

THEOREM 3.2. *For CG approximations of degree n, with $1 < n \le 20$, the dispersion relation obtained from (3.12) is the polynomial of degree n:*

$$P_n^{CG}(\widetilde{\omega}) = \widetilde{\omega}^n + \sum_{j=1}^{\lfloor \frac{n+1}{2} \rfloor} \widetilde{a}_{n-2j+1}\, \widetilde{\omega}^{n-2j+1} + \sum_{j=1}^{\lfloor \frac{n}{2} \rfloor} \widetilde{a}_{n-2j}\, \widetilde{\omega}^{n-2j} = 0, \tag{3.13}$$

where floor$(x) = \lfloor x \rfloor$ and $\binom{q}{r} = \frac{q!}{(q-r)!r!}$, with

$$a_{n-2j+1} = \frac{(-1)^j}{n^{2j-1}} \binom{n}{2j-1} \frac{(n+2j)!}{(n+1)!}, \quad a_{n-2j} = \frac{(-1)^j}{n^{2j}} \binom{n}{2j} \frac{(n+2j+1)!}{(n+1)!}, \tag{3.14}$$

$$\frac{\widetilde{a}_{n-2j+1}}{a_{n-2j+1}} = \frac{\sin(nkh)}{\cos(nkh) - (-1)^n(n+1)}, \quad \frac{\widetilde{a}_{n-2j}}{a_{n-2j}} = \frac{\cos(nkh) - (-1)^n\left(\frac{n+1}{2j+1}\right)}{\cos(nkh) - (-1)^n(n+1)}. \tag{3.15}$$

These results have been obtained using a computer algebra system (Maple). It is conjectured that (3.13)–(3.15) also hold for all $n > 20$. Let $\widetilde{a}_n = a_n := 1$, and note that the coefficient $\widetilde{a}_j$, $j = 1, 2, \ldots$, has two different meanings depending on whether n is even or odd.

The case $n = 1$ yields $\widetilde{\omega}\big(\cos(kh) + 2\big) = 3\sin(kh)$, and $\widetilde{\omega}$ is graphed in Figure 4 with the continuous frequency $\widetilde{\omega}^{AN} := kh$. The observed discrepancy is discussed later.

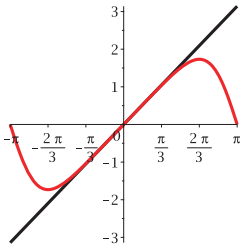

FIG. 4. *The continuous frequency $\widetilde{\omega}^{AN}$ (black) and the computed frequency $\widetilde{\omega}$ (red) for the CG scheme in the case $n = 1$.*

COROLLARY 1. *Several symmetries for $\widetilde{\omega}(kh)$ result from (3.13)–(3.15). First, $\widetilde{\omega}_j(kh) = \widetilde{\omega}_j\big(kh + \frac{2\pi}{n}\big)$, $j = 1, 2, \ldots, n$. Second, let $p = 1, 2, 3, \ldots$ and $j = 1, 2, \ldots, p$. For $n = 2p + 1$, $2p$ solutions satisfy $\widetilde{\omega}_j(kh) = -\widetilde{\omega}_{2(p+1)-j}(-kh)$ and one solution reads as $\widetilde{\omega}_{p+1}(kh) = -\widetilde{\omega}_{p+1}(-kh)$, and only $\widetilde{\omega}_{p+1}$ is zero at $kh = 0$. For $n = 2p$, $2p$ solutions satisfy $\widetilde{\omega}_j(kh) = -\widetilde{\omega}_{2p+1-j}(-kh)$ and both $\widetilde{\omega}_p$ and $\widetilde{\omega}_{p+1}$ vanish at $kh = 0$.*

COROLLARY 2. *In the limit as mesh spacing $kh \to 0$, the group velocities $\frac{\partial \widetilde{\omega}_j}{\partial kh}$, $j = 1, 2, \ldots, n$, obtained from the solutions $\widetilde{\omega}_j$ of $P_n^{CG}(\widetilde{\omega})$ in (3.13) satisfy the following:*

The case $n = 2p + 1$, $p = 1, 2, 3, \ldots$	*The case $n = 2p$, $p = 1, 2, 3, \ldots$*
For $j = 1, 2, \ldots, n$, $\lim_{kh\to 0} \frac{\partial \widetilde{\omega}_j}{\partial kh} = 1$.	*For $j = 1, 2, \ldots, p-1, p+2, \ldots, n$,* $\lim_{kh\to 0} \frac{\partial \widetilde{\omega}_j}{\partial kh} = 1$, *and for $j = p$ and $j = p+1$,* $\lim_{kh\to 0} \frac{\partial \widetilde{\omega}_j}{\partial kh} = (-1)^j(n+1) - n$; *in other words, $-(2n+1)$ or 1 depending on whether p is even or odd.*

Proof. Since $\frac{dP_n^{CG}}{d\,kh} = \frac{\partial P_n^{CG}}{\partial kh} + \frac{\partial P_n^{CG}}{\partial \widetilde{\omega}}\,\frac{\partial \widetilde{\omega}}{\partial kh} = 0$, (3.13) yields

$$\frac{dP_n^{CG}}{d\,kh} = \sum_{j=1}^{\lfloor \frac{n+1}{2} \rfloor} \frac{\partial \widetilde{a}_{n-2j+1}}{\partial kh}\,\widetilde{\omega}^{n-2j+1} + \sum_{j=1}^{\lfloor \frac{n}{2} \rfloor} \frac{\partial \widetilde{a}_{n-2j}}{\partial kh}\,\widetilde{\omega}^{n-2j} + \frac{\partial \widetilde{\omega}}{\partial kh}\left(\sum_{j=1}^{\lfloor \frac{n}{2} \rfloor} (n-2j+1)\widetilde{a}_{n-2j+1}\,\widetilde{\omega}^{n-2j} + \sum_{j=1}^{\lfloor \frac{n+1}{2} \rfloor} (n-2j+2)\widetilde{a}_{n-2j+2}\,\widetilde{\omega}^{n-2j+1} \right) = 0. \tag{3.16}$$

We obtain $\lim_{kh\to 0} \frac{\partial \widetilde{\omega}}{\partial kh} = \lim_{kh\to 0} \frac{\Psi_1}{\Psi_2}$ from (3.16) by letting

$$\Psi_1 = \sum_{j=1}^{\lfloor \frac{n+1}{2} \rfloor} \frac{(n-2j+2)(n+2j)}{(2j-1)}\,a_{n-2j+2}\,\widetilde{\omega}^{n-2j+1}, \tag{3.17}$$

$$\Psi_2 = \sum_{j=1}^{\lfloor \frac{n+1}{2} \rfloor} \frac{(n-2j+2)\Big((-1)^{n+1}(n+1)+2j-1\Big)}{(2j-1)}\,a_{n-2j+2}\,\widetilde{\omega}^{n-2j+1}. \tag{3.18}$$

In the case of n odd, with $n = 2p+1$, $p = 1, 2, 3, \ldots$, we deduce $\Psi_1 = \Psi_2$ from (3.17)–(3.18) with $a_1 \neq 0$, and consequently $\lim_{kh\to 0} \frac{\partial \widetilde{\omega}}{\partial kh} = 1$.

In the case of n even, with $n = 2p$, $p = 1, 2, 3, \ldots$, we obtain from (3.17)–(3.18)

$$\Psi_1 - \Psi_2 = 2(n+1) \sum_{j=1}^{p} \frac{(n-2j+2)}{(2j-1)}\,a_{n-2j+2}\,\widetilde{\omega}^{n-2j+1}, \tag{3.19}$$

and (3.13)–(3.15) lead to

$$0 = \lim_{kh\to 0} P_n^{CG}(\widetilde{\omega}) = \lim_{kh\to 0} \left(\widetilde{\omega}^n + \frac{1}{n} \sum_{j=1}^{p} \frac{(n-2j)}{(2j+1)}\,a_{n-2j}\,\widetilde{\omega}^{n-2j} \right). \tag{3.20}$$

The lowest degree term with respect to $\widetilde{\omega}$ in the right-hand side (RHS) of (3.20) is $\frac{2}{n(n-1)}\,a_2\,\widetilde{\omega}^2$, with $a_2 \neq 0$, and two cases need to be considered: $\widetilde{\omega} \neq 0$ and $\widetilde{\omega} = 0$.

When $\widetilde{\omega} \neq 0$, dividing (3.20) by $\widetilde{\omega}$ and using (3.19) yields $\lim_{h\to 0} (\Psi_1 - \Psi_2) = 0$. Since $a_2 \neq 0$, we have $\lim_{kh\to 0} \Psi_2 \neq 0$ from (3.18), and hence $\lim_{h\to 0} \frac{\partial \widetilde{\omega}_j}{\partial kh} = 1$ for $j = 1, 2, \ldots, p-1, p+2, \ldots, 2p$, i.e., the nonzero solutions $\widetilde{\omega}_j$ at $kh = 0$ when n is even.

When $\widetilde{\omega} = 0$, we have $a_0 = 0$ and $a_1 = 0$ at $kh = 0$ from (3.14) and (3.15). In order to obtain the group velocities for $\widetilde{\omega}_p$ and $\widetilde{\omega}_{p+1}$, we thus need to consider

$$\frac{d^2 P_n^{CG}}{d\,(kh)^2} = \frac{\partial^2 P_n^{CG}}{\partial \widetilde{\omega}^2} \left(\frac{\partial \widetilde{\omega}}{\partial kh} \right)^2 + 2\,\frac{\partial^2 P_n^{CG}}{\partial \widetilde{\omega} \partial kh} \left(\frac{\partial \widetilde{\omega}}{\partial kh} \right) + \frac{\partial^2 P_n^{CG}}{\partial (kh)^2} + \frac{\partial P_n^{CG}}{\partial \widetilde{\omega}}\,\frac{\partial^2 \widetilde{\omega}}{\partial (kh)^2} = 0$$

and compute $\lim_{kh\to 0} \frac{\partial \widetilde{\omega}_{p,p+1}}{\partial kh} = \lim_{kh\to 0} \frac{1}{2a_2} \big(- \frac{\partial a_1}{\partial kh} \pm \big(\big(\frac{\partial a_1}{\partial kh} \big)^2 - 2a_2 \frac{\partial^2 a_0}{\partial (kh)^2} \big)^{1/2} \big)$. ∎

3.3. The presence of erratic modes.

THEOREM 3.3. *The following properties hold:*

(i) *For the matrix* $\mathcal{G}$ *of size* $(n+1)\times(n+1)$ *defined in* (3.4)*, we have* $\operatorname{rank}(\mathcal{G}) = n$.

(ii) *A leading principal submatrix of $\mathcal{G}$ of size $n \times n$, denoted by $\mathcal{G}^\diamond$ and obtained by removing the first column and the last row of $\mathcal{G}$, is nonsingular, with*

$$\mathcal{G}^\diamond = (G_{q,r}) \quad \textit{for } 0 \le q \le n-1 \textit{ and } 1 \le r \le n. \tag{3.21}$$

(iii) *For all $n \in \mathbb{N}$, the CG scheme admits a unique erratic stationary mode.*

Proof. We first prove property (i). A vector $(\zeta_1, \zeta_2, \zeta_3, \dots, \zeta_{n+1})^T$ is in the kernel or the null space of $\mathcal{G}$, denoted by $\ker(\mathcal{G})$, if and only if, for all r such that $0 \le r \le n$, the polynomial $\sum_{q=0}^{n} \zeta_{q+1}\, v_{j+\frac{q}{n}}$ on e_j belongs to the orthogonal of the space Υ, denoted by $\Upsilon^\perp$, generated by the vectors $\big(\frac{\partial v_{j+\frac{r}{n}}}{\partial x}\big)\Big|_{0\le r\le n}$, for the scalar product on $\mathbb{R}_n[x]$ defined as $< v_k, v_l >= \int_{e_j} v_k v_l\, dx$. Since $\big(v_{j+\frac{q}{n}}\big)_{0\le q\le n}$ is a basis of $\mathbb{R}_n[x]$, we have $\dim\big(\ker(\mathcal{G})\big) = \dim(\Upsilon^\perp)$ in $\mathbb{R}_n[x]$, and thus we need to show that $\dim(\Upsilon) = n$. Let Θ be a linear mapping such that for $\Xi = (\xi_0, \xi_1, \dots, \xi_n)^T$ belonging to $\mathbb{R}^{n+1}$ we have $\Theta(\Xi) = \sum_{r=0}^{n} \xi_r \frac{\partial v_{j+\frac{r}{n}}}{\partial x}$. Then, Ξ belongs to the kernel of Θ if and only if $\sum_{r=0}^{n} \xi_r v_{j+\frac{r}{n}}$ is a constant. The fact that $\big(v_{j+\frac{q}{n}}\big)_{0\le q\le n}$ is a basis of $\mathbb{R}_n[x]$ ends the proof.

We now prove (ii). Property (i) yields $\dim\ker(\mathcal{G}) = \dim\ker(\mathcal{G}^T) = 1$. Let $\mathbf{u}^{\ker\mathcal{G}}$ and $\mathbf{u}^{\ker\mathcal{G}^T}$ be the vectors with $n+1$ components belonging to $\ker(\mathcal{G})$ and $\ker(\mathcal{G}^T)$, respectively. On element e_j of ε_h, $j = 1, 2, \dots, m$, we have

$$\mathbf{u}^{\ker\mathcal{G}} = (1, u_1^{\ker\mathcal{G}}, \dots, u_{n-1}^{\ker\mathcal{G}}, u_n^{\ker\mathcal{G}})^T, \tag{3.22}$$

$$\mathbf{u}^{\ker\mathcal{G}^T} = (u_0^{\ker\mathcal{G}^T}, u_1^{\ker\mathcal{G}^T}, \dots, u_{n-1}^{\ker\mathcal{G}^T}, 1)^T, \tag{3.23}$$

where $\mathbf{u}^{\ker\mathcal{G}}$ and $\mathbf{u}^{\ker\mathcal{G}^T}$ are normalized by setting $u_0^{\ker\mathcal{G}} = 1$ and $u_n^{\ker\mathcal{G}^T} = 1$, respectively. Further, let $\mathcal{G}^L$ and $\mathcal{G}^R$ be the matrices of size $(n+1)\times n$ and $n \times (n+1)$,

(3.24)

$$\mathcal{G}^L = \begin{pmatrix} & & \mathcal{I}_{n\times n} & & \\ & & & & \\ -u_0^{\ker\mathcal{G}^T} & -u_1^{\ker\mathcal{G}^T} & \cdots & -u_{n-2}^{\ker\mathcal{G}^T} & -u_{n-1}^{\ker\mathcal{G}^T} \end{pmatrix},\ \mathcal{G}^R = \begin{pmatrix} -u_1^{\ker\mathcal{G}} & \\ -u_2^{\ker\mathcal{G}} & \\ & \mathcal{I}_{n\times n} \\ -u_{n-1}^{\ker\mathcal{G}} & \\ -u_n^{\ker\mathcal{G}} & \end{pmatrix},$$

respectively, where $\mathcal{I}_{n\times n}$ is the $n \times n$ identity matrix.

The computation of $\mathcal{G}^\diamond\,\mathcal{G}^R$ yields an $n \times (n+1)$ matrix coinciding with the n first lines of $\mathcal{G}$ due to (3.6). Multiplying $\mathcal{G}^\diamond\,\mathcal{G}^R$ on the left by $\mathcal{G}^L$ yields $\mathcal{G}^L\,\mathcal{G}^\diamond\,\mathcal{G}^R = \mathcal{G}$ since $\mathcal{G}^T\mathbf{u}^{\ker\mathcal{G}^T} = \mathbf{0}$ and $u_n^{\ker\mathcal{G}^T} = 1$. Further, $\operatorname{rank}(\mathcal{G}^L) = \operatorname{rank}(\mathcal{G}^R) = n$ from (3.24) and $\operatorname{rank}(\mathcal{G}) = n$ from (i). Since $\operatorname{rank}(\mathcal{G}^L\,\mathcal{G}^\diamond\,\mathcal{G}^R) \le \min(n, \operatorname{rank}(\mathcal{G}^\diamond))$, we deduced from $\mathcal{G} = \mathcal{G}^L\,\mathcal{G}^\diamond\,\mathcal{G}^R$ that $\operatorname{rank}(\mathcal{G}^\diamond) = n$; then $\mathcal{G}^\diamond$ is full rank with $\det(\mathcal{G}^\diamond) \neq 0$.

Finally, we prove property (iii). Stationary erratic modes occur when $\omega = 0$ in (3.2), and the kernel of the discrete gradient operator in (3.8)–(3.9) is thus computed. For symmetry reasons, selected discrete equations are considered, namely

(i) $n-1$ equations at interior nodes $j + \frac{q}{n}$ over e_j,
(ii) $n-1$ equations at interior nodes $j - \frac{q}{n}$ over e_{j-1} for $1 \le q \le n-1$,
(iii) one equation at a typical boundary element node j.

From (3.8)–(3.9) a system of $2n-1$ equations is then obtained with the $2n+1$

unknowns $u_{j-1}, u_{j-\frac{n-1}{n}}, \ldots, u_{j-\frac{1}{n}}, u_j, u_{j+\frac{1}{n}}, \ldots, u_{j+\frac{n-1}{n}}, u_{j+1}$,

$$\pm \sum_{r=0}^{n} u_{j\pm\frac{r}{n}} G_{q,r} = 0 \quad \text{for } 1 \le q \le n-1, \tag{3.25}$$

$$\sum_{r=0}^{n} \left(u_{j+\frac{r}{n}} - u_{j-\frac{r}{n}} \right) G_{0,r} = 0. \tag{3.26}$$

By adding the equations in (3.25), and using (3.26), a system of n equations is obtained:

$$\sum_{r=1}^{n} \left(u_{j+\frac{r}{n}} - u_{j-\frac{r}{n}} \right) G_{q,r} = 0 \quad \text{for } 0 \le q \le n-1. \tag{3.27}$$

The matrix of the linear system (3.27) is nothing more than $\mathcal{G}^{\diamond}$, with $\det(\mathcal{G}^{\diamond}) \neq 0$ obtained from (ii). Hence, (3.27) yields $u_{j+\frac{r}{n}} = u_{j-\frac{r}{n}}$ for $1 \le r \le n$.

It remains to solve (3.25), a system of $n-1$ equations with the $n+1$ unknowns $u_j, u_{j+\frac{1}{n}}, u_{j+\frac{2}{n}}, \ldots, u_{j+\frac{n-1}{n}}, u_{j+1}$. Let $\mathcal{G}^{\square}$ be the $(n-1) \times (n+1)$ matrix in the left-hand side (LHS) of (3.25), obtained by removing the first and the last rows of $\mathcal{G}$, namely

$$\mathcal{G}^{\square} = (G_{q,r}) \quad \text{for } 1 \le q \le n-1 \text{ and } 0 \le r \le n. \tag{3.28}$$

Further, let $\mathbf{g}_r^{\square}$ be the $n+1$ column vectors of $\mathcal{G}^{\square}$ for $0 \le r \le n$ and $\mu_r \in \mathbb{R}$ for $0 \le r \le n$. Due to (3.6), we obtain

$$\sum_{r=0}^{n} \mu_r \, \mathbf{g}_r^{\square} = \sum_{r=1}^{n} (\mu_r - \mu_0) \, \mathbf{g}_r^{\square}. \tag{3.29}$$

Since the n column vectors in the RHS of (3.29) coincide with the $n-1$ row vectors of $\mathcal{G}^{\diamond}$, with $\operatorname{rank}(\mathcal{G}^{\diamond}) = n$, we deduce that $\operatorname{rank}(\mathcal{G}^{\square}) = n-1$ and, consequently, that $\dim \ker(\mathcal{G}^{\square}) = 2$. Two independent vectors thus lie in the kernel of $\mathcal{G}^{\square}$. The first vector is the geostrophic mode of the form $(1, 1, 1, \ldots, 1)^T$, which reflects the fact that u is defined up to a constant in (2.1). It corresponds to $\mathbf{u}^{\ker \mathcal{G}}$ in (3.22), and it is computed by using (3.6). The second vector takes the form of a unique erratic stationary mode. It is the highest frequency mode made up of alternating positive and negative values and has the largest phase speed error. It corresponds to $\mathbf{u}^{\ker \mathcal{G}^T}$ in (3.23), and it is computed from (3.25) over elements e_{j-1} and e_j, $j = 1, 2, \ldots, m$. In Figure 2 and for $n = 1, 2, 3$, we obtain $\mathbf{u}^{\ker \mathcal{G}^T} = (1, -1, 1)$, $\mathbf{u}^{\ker \mathcal{G}^T} = (1, -1/2, 1, -1/2, 1)$, and $\mathbf{u}^{\ker \mathcal{G}^T} = (1, -11/27, 11/27, -1, 11/27, -11/27, 1)$, respectively. ∎

3.4. The presence of spectral gaps. Rearranging terms in (3.13) yields

$$Q_n(\widetilde{\omega}) \cos(nkh) + R_{n-1}(\widetilde{\omega}) \sin(nkh) + S_n(\widetilde{\omega}) = 0, \tag{3.30}$$

where Q_n, R_{n-1}, and S_n are polynomials of degree n or $n-1$ with respect to $\widetilde{\omega}$, with

$$Q_n(\widetilde{\omega}) = \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} a_{n-2j} \, \widetilde{\omega}^{n-2j}, \quad R_{n-1}(\widetilde{\omega}) = \sum_{j=1}^{\lfloor \frac{n+1}{2} \rfloor} a_{n-2j+1} \, \widetilde{\omega}^{n-2j+1},$$

$$S_n(\widetilde{\omega}) = (-1)^{n+1}(n+1) \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} \frac{a_{n-2j}}{2j+1} \, \widetilde{\omega}^{n-2j}.$$

COROLLARY 3. *The roots* $\widetilde{\omega}_j$, $j = 1, 2, \ldots, n$, *of* $P_n^{CG}(\widetilde{\omega})$ *in* (3.13) *do not exist if*

$$|S_n(\widetilde{\omega})| > \sqrt{Q_n^2(\widetilde{\omega}) + R_{n-1}^2(\widetilde{\omega})}. \tag{3.31}$$

Proof. After long and tedious algebra we obtain for all n

$$Q_n^2(\widetilde{\omega}) + R_{n-1}^2(\widetilde{\omega}) = \sum_{j=0}^{n} \binom{2j+1}{j} \frac{(n+j+1)!}{n^{2j}(n+1)(n-j)!}\, \widetilde{\omega}^{2(n-j)}, \tag{3.32}$$

and hence $Q_n^2(\widetilde{\omega}) + R_{n-1}^2(\widetilde{\omega}) > 0$ since the lowest degree term in (3.32) for $j = n$, with respect to $\widetilde{\omega}$, is different from zero. Consequently, the proof ends by letting $(\sin\varphi, \cos\varphi) = \left(Q_n^2(\widetilde{\omega}) + R_{n-1}^2(\widetilde{\omega})\right)^{-1/2} (Q_n(\widetilde{\omega}), R_{n-1}(\widetilde{\omega}))$ and rewriting (3.30) as $S_n(\widetilde{\omega}) = -\sqrt{Q_n^2(\widetilde{\omega}) + R_{n-1}^2(\widetilde{\omega})}\, \sin(nkh + \varphi)$. ∎

When (3.31) is satisfied, $\widetilde{\omega}$ is thus no longer continuous in kh and it exhibits discontinuities or gaps. A few intervals then exist, given for $n = 3, 4, \ldots, 8$ in Table 1, where $\widetilde{\omega}$ is not defined. The gaps have been observed in [12], but this is the first time, to our knowledge, that their existence and exact location are proven analytically.

TABLE 1
Intervals of values of $\widetilde{\omega}$, *in the case* $n = 1, 2, \ldots, 8$, *for which the condition* (3.31) *is satisfied, namely when* $\widetilde{\omega}$ *cannot be computed. There is no gap when* $n = 1, 2$.

n	$\mathcal{I}_{\widetilde{\omega}}(n)$
3	[0.8820 , 0.9481]
4	[1.323 , 1.525]
5	[0.57463 , 0.57574] ∪ [1.597 , 1.952]
6	[0.96357 , 0.97461] ∪ [1.794 , 2.311]
7	[0.42053 , 0.42054] ∪ [1.239 , 1.273] ∪ [1.953 , 2.641]
8	[0.74050 , 0.74085] ∪ [1.442 , 1.512] ∪ [2.094 , 2.959]

3.5. The graphs of the dispersion relation. The n solutions of $P_n^{CG}(\widetilde{\omega})$ are shown in Figures 5 to 6 for $n = 2, \ldots, 7$, on the first line, distinguishing between even (Figure 5) and odd n (Figure 6). At a given spatial wavenumber kh there are n solutions, and each solution is given its own color. Each of these n solutions represents the n Fourier modes present in each eigenmode solution of the discrete equation. This fact has caused some confusion in the literature, with several works associating the single solution with n different solutions, one physical and $n-1$ unphysical spurious modes [6, 8]. This interpretation results in a multiple-valued dispersion relation $\widetilde{\omega}(kh)$. We find this interpretation misleading and prefer the approach taken in [12, 14, 15], where each eigenmode is associated with a single kh rather than treated as n different solutions with n different values of kh. It is also more consistent with the Bloch wave approach, where the phase velocity is a single-valued function of each Bloch mode. However, for $n \geq 4$, the maximum amplitude approach employed in [12] leads to a dispersion relation for high kh that is multiple valued with large gaps near $4h$ [12, Figure 11], making it challenging to compute the group velocity at high frequencies.

Here a modified kh identification procedure is proposed which consists in considering each of the n solutions of Figures 5 to 6 (first line) valid over only a limited wavenumber range, termed a branch. The union of all branches then gives the

complete dispersion relationship, named $\widetilde{\omega}_S$, in Figures 5 to 6 (second line). All the remaining solutions are mathematical artifacts arising from symmetries in the Fourier analysis (see Corollary 1). The physical solution and the mathematical artifacts can be distinguished by inspecting the spatial structure of u_h for each branch, using the eigenvectors. The spatial structure is computed as follows: First, numerically solve (3.12) for a given kh, yielding a set of n eigenvalues and eigenvectors, which are the n Fourier amplitudes. The solution can then be reconstructed using the definition of nodal values in terms of Fourier amplitudes. In Figures 7 to 10, a mesh with 12 elements of unit width is chosen and $h = 1/n$. Examples are given for several values of n and kh.

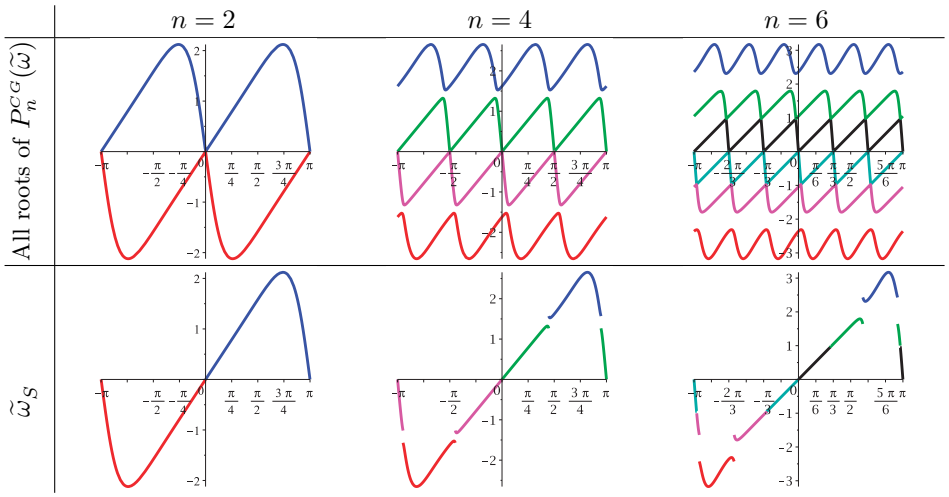

FIG. 5. *Phase $\widetilde{\omega}$ for the CG scheme in the case $n = 2, 4, 6$. For a given n, the first line corresponds to all solutions of $P_n^{CG}(\widetilde{\omega})$ and $\widetilde{\omega}_S$ is given in the second line.*

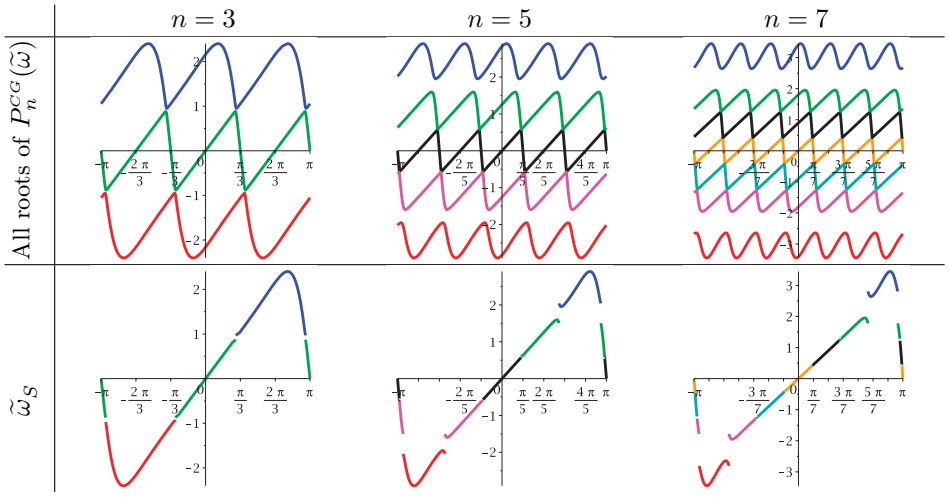

FIG. 6. *As for Figure 5 but in the case $n = 3, 5, 7$.*

For $n = 4$ and $kh = \pi/4$, the reconstructed solutions are displayed in Figure 7 according to the colors of Figure 5. The red curve corresponds to $kh = -3\pi/4$, not

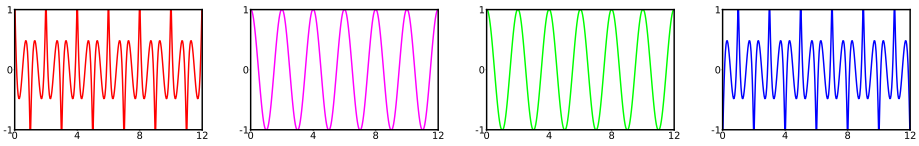

FIG. 7. *Reconstructed solutions for the CG method with $n = 4$ at $kh = \pi/4$. The colors correspond with the associated branches in Figure 5.*

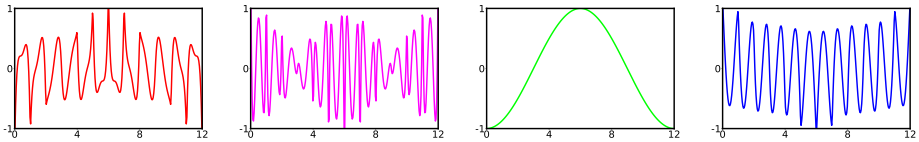

FIG. 8. *As for Figure 7 but in the case $n = 4$ at $kh = \pi/24$.*

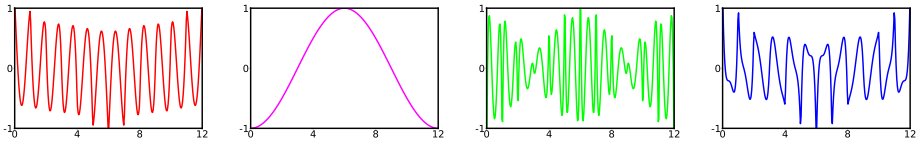

FIG. 9. *As for Figure 7 but in the case $n = 4$ at $kh = -\pi/24$.*

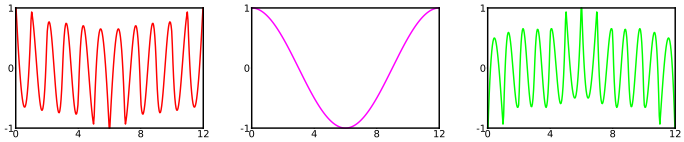

FIG. 10. *As for Figure 7 but in the case $n = 3$ at $kh = \pi/18$. The colors correspond with the associated branches in Figure 6.*

$kh = \pi/4$, while the magenta curve corresponds to $kh = -\pi/4$, not $kh = \pi/4$. The green curve corresponds to $kh = -\pi/4$, as expected, while the blue curve corresponds to $kh = 3\pi/4$, not $kh = \pi/4$. Clearly, the reconstructed solutions with negative kh represent a mirror symmetry. Identical results are obtained if we instead compute the eigenvectors associated with $kh = -3\pi/4$, $kh = -\pi/4$, or $kh = 3\pi/4$ (not shown). The reconstructed solutions are also shown in Figures 8 to 9 at high wavelength for $kh = \pm\pi/24$, and we observe that considering the two figures, the green and magenta curves exchange and the red and blue curves also exchange. The stationary mode, in the form of a wave packet consisting of a superposition of a high-frequency mode near the grid scale and a low-frequency mode near $kh = 0$, is clearly shown in the magenta solution in Figure 8 at $kh = \pi/24$ and the green solution at $kh = -\pi/24$ in Figure 9. For $n = 3$ and $kh = \pi/18$, the reconstructed solutions are shown in Figure 10. Three identical solutions are obtained for $kh = -\pi/18$, and they are not displayed. These results clearly illustrate the consequences of Corollary 2, depending on whether n is even or odd.

Although the above results are not a formal proof, they are a strong argument for the validity of the branch selection procedure used to construct $\widetilde{\omega}_S$ from $P_n^{CG}(\widetilde{\omega})$, shown in the second line of Figures 5 to 6. Indeed, even in the high wavenumber part of the dispersion relationship, where the solutions are quite erratic and have minimal correspondence to the analytical solutions, the group velocity correctly predicts the

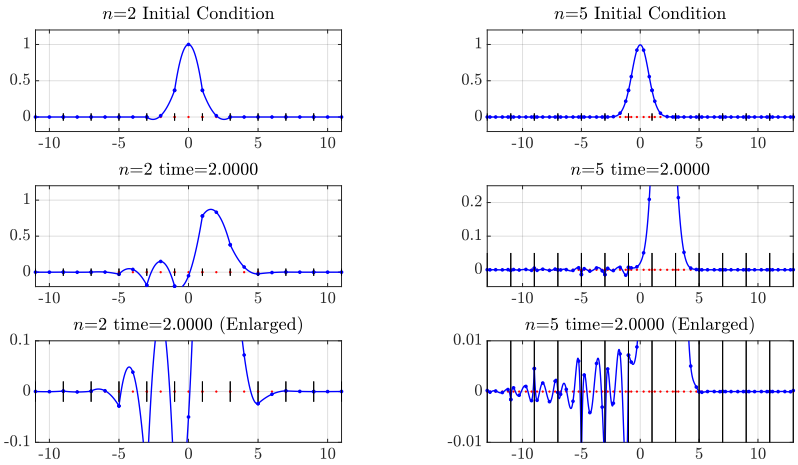

FIG. 11. *Propagation of an initial Gaussian on a mesh of* 11 *elements for* $n = 2$ *and* 5. *An anomalous wave packet with a predicted group velocity of* $-(2n+1)$ *is visible in the enlarged figures.*

wave packet behavior. With this kh identification procedure, there is a one-to-one correspondence between each discrete eigenmode and an associated propagating wave solution to the continuum equations. Thus, the class of methods analyzed here does not contain any spurious modes if spurious modes are defined as eigenmodes representing spurious numerical solutions with no corresponding solution in the continuum equations. However, many of the higher frequency eigenmodes have spuriously negative group velocities, and this is why we refer to these modes as erratic. In particular, as with many collocated discretizations, the highest frequency eigenmode is in the null space of the derivative operator and is thus stationary (see subsection 3.3).

The spectral gaps discussed in Corollary 3, which occur for $n \geq 3$, are clearly visible in Figures 5 to 6. Wave packets with energy at this wavenumber fail to propagate correctly, leading to significant numerical dispersion. The gaps always come in pairs, and there are floor$((n-1)/2)$ pairs. Although the number of pairs increases as n increases, the gaps in the low-frequency part of the spectrum decrease in size, while only the last pair in the highest-frequency part of the spectrum increases.

Finally, consider the behavior at the end of the spectrum. For both even and odd n, the slope of the dispersion relationship is obtained similarly to the proof of Corollary 2 and at $kh = \pi$ it is $-(2n+1)$ with $\widetilde{\omega} = 0$ (corresponding to the erratic stationary mode). Such an incorrect slope is key to understanding the behavior of discrete simulations. Consider the propagation of a Gaussian as shown in Figure 11. A mesh of 11 elements of width 2 is considered with $n = 2, 3, 4, 5$, but only the results for $n = 2, 5$ are shown. In all cases, there is an anomalous wave packet with a group velocity of $-(2n+1)$ (as predicted). This can be understood as follows. The Gaussian initial condition projects onto a range of wavenumbers, and for a given $\widetilde{\omega}$ (when it is representable), there are two spatial wavenumbers, giving rise to two distinct wave packets. One wave packet will have a group velocity close to 1 and the other group velocity close to $-(2n+1)$. Both of these wave packets are clearly visible in Figure 11.

Additionally, the maximal frequency increases as n increases, which is a likely cause of the observation that the CFL limit gets progressively stricter with higher n. In the next section, we now consider the discontinuous approximation of (2.2).

4. Discontinuous Galerkin discretizations.

4.1. Discontinuous Galerkin formulations. The setting of subsection 3.1 is still used, except that ε_h is now a finite collection of m *open* elements e_j, $j = 1, 2, \ldots, m$, of the real line, such that $\bar{\Omega} = \bigcup_{e_j \in \varepsilon_h} \bar{e}_j$ and $e_i \cap e_j = \emptyset$ for $i \neq j$. The so-called *broken space* $H^1(\varepsilon_h)$ is defined as $H^1(\varepsilon_h) = \{v \in L^2(\Omega);\ v|_e \in H^1(e)$ for all $e \in \varepsilon_h\}$.

Let u be a sufficiently smooth function. Multiplying (2.2) by a function v belonging to $H^1(\varepsilon_h)$ and integrating over the domain Ω leads to

$$i\omega \sum_{j=1}^{m} \int_{e_j} u\, v\, dx - c \sum_{j=1}^{m} \int_{e_j} \frac{\partial u}{\partial x}\, v\, dx = 0, \tag{4.1}$$

an equation similar to (3.1), except that u and v are now discontinuous at the element boundaries of e_j, $j = 1, 2, \ldots, m$, and the integrals are computed over e_j and not Ω. To obtain the DG formulation, the integrals in (4.1) are integrated by parts, yielding

$$i\omega \sum_{j=1}^{m} \int_{e_j} u\, v\, dx + c \sum_{j=1}^{m} \left(\int_{e_j} u\, \frac{\partial v}{\partial x}\, dx - u^* v \,|_{j^+}^{(j+1)^-} \right) = 0, \tag{4.2}$$

where u^* denotes the numerical trace of u, and j^- and j^+ are the nodal positions of adjacent elements corresponding to a typical node j. That is, node j corresponds to the coincident node pair (j^-/j^+), $j = 1, 2, 3, \ldots, m+1$, as shown in Figure 12. Since u^* is uniquely defined at the element boundary, we obtain

$$\sum_{j=1}^{m} u^* v \,|_{j^+}^{(j+1)^-} = \sum_{j=1}^{m} \left(u^*_{j+1} v(x_{(j+1)^-}) - u^*_j v(x_{j^+}) \right) = \sum_{j=1}^{m} u^*_j \big(v(x_{j^-}) - v(x_{j^+}) \big), \tag{4.3}$$

by developing the sum and applying the periodic boundary condition.

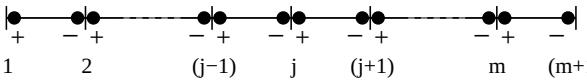

FIG. 12. *Node j corresponds to the coincident node pair (j^-/j^+), $j = 1, 2, 3, \ldots, m+1$.*

Let $[\,u(x_j)\,] = u(x_{j^-}) - u(x_{j^+})$ and $\{u(x_j)\} = \frac{1}{2}\big(u(x_{j^-}) + u(x_{j^+})\big)$ be the jump and the mean of u at node j, respectively. The numerical trace is chosen as

$$u^*_j := u^*_j(\lambda) = \{u(x_j)\} + \left(\frac{1}{2} - \lambda\right)[\,u(x_j)\,] = (1-\lambda)\, u(x_{j^-}) + \lambda\, u(x_{j^+}), \tag{4.4}$$

where λ is a real parameter leading to the upwind and centered fluxes for $\lambda = 1$ and $\lambda = 1/2$, respectively. The variational formulation for the DG method then reads as

$$i\omega \sum_{j=1}^{m} \int_{e_j} u\, v\, dx + c \sum_{j=1}^{m} \int_{e_j} u\, \frac{\partial v}{\partial x}\, dx - c \sum_{j=1}^{m} u^*_j(\lambda)\, [\,v(x_j)\,] = 0. \tag{4.5}$$

In the DG approximation, we consider finite element spaces V_h^n of polynomial functions, discontinuous at the element interfaces, such that

$$V_h^n = \{\, u_h \in L^2(\Omega);\ u_h|_e = \widetilde{u}_h \circ F_e^{-1},\ \widetilde{u}_h \in \mathcal{P}_n(\widetilde{e})\ \forall\, e \in \varepsilon_h\}.$$

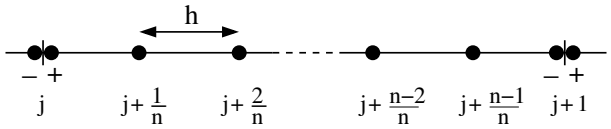

FIG. 13. *Indices of the local degrees of freedom on element e_j of ε_h, $j = 1, 2, ..., m$.*

The basis $\mathcal{V}_h^n$ of V_h^n is a basis of the $n+1$ Lagrange polynomials of degree n, with $\mathcal{V}_h^n = \{v_s\}$, for $s \in \mathcal{J}_j^n$, where $\mathcal{J}_j^n = \{j^+, j+\frac{1}{n}, j+\frac{2}{n}, j+\frac{3}{n}, \ldots, j+\frac{n-1}{n}, (j+1)^-\}$, $j = 1, 2, \ldots, m$, is the set of indices of the local degrees of freedom on each element e_j of ε_h, shown in Figure 13. Further, h is defined as in subsection 3.1. Introducing the basis $\mathcal{V}_h^n$ leads to a discretization of (4.5), which consists in finding u_h belonging to V_h^n for the selected bases, at node $s \in \mathcal{J}_j^n$, such that

$$i\omega \sum_{j=1}^{m} \int_{e_j} u_h v_s \, dx + c \sum_{j=1}^{m} \int_{e_j} u_h \frac{\partial v_s}{\partial x} \, dx - c \sum_{j=1}^{m} u_j^*(\lambda) \, [\, v_s(x_j) \,] = 0 \quad \forall v_s \in V_h^n. \tag{4.6}$$

The computation of (4.6) over e_j at nodes $s = j + \frac{q}{n}$, for $q = 1, 2, \ldots, n-1$, $s = j^+$, and $s = (j+1)^-$, $j = 1, 2, \ldots, m$, is now performed for the purpose of the Fourier analysis. At $s = j + \frac{q}{n}$, $q = 1, 2, \ldots, n-1$, we have $[\, v_s(x_j) \,] = 0$, $j = 1, 2, \ldots, m$, and we obtain

$$\begin{aligned} &u_{j^+}\big(i\widetilde{\omega} M_{q,0} - G_{q,0}\big) + \sum_{r=1}^{n-1} u_{j+\frac{r}{n}}\big(i\widetilde{\omega} M_{q,r} - G_{q,r}\big) \\ &+ u_{(j+1)^-}\big(i\widetilde{\omega} M_{q,n} - G_{q,n}\big) = 0. \end{aligned} \tag{4.7}$$

Since we are only concerned with element e_j, there is no ambiguity to set the indices of elements of $\mathcal{M}$ and $\mathcal{G}$ in (3.4) corresponding to nodes j^+ and $(j+1)^-$ to 0 and n, respectively. Evaluating (4.6) at $s = j^+$, $j = 1, 2, \ldots, m$, where $[\, v_{j^+}(x_j) \,] = -1$, yields

$$\begin{aligned} &u_{j^-}(1-\lambda) + u_{j^+}\big(i\widetilde{\omega} M_{0,0} + G_{0,0} + \lambda\big) + \sum_{r=1}^{n-1} u_{j+\frac{r}{n}}\big(i\widetilde{\omega} M_{0,r} - G_{0,r}\big) \\ &+ u_{(j+1)^-}\big(i\widetilde{\omega} M_{0,n} - G_{0,n}\big) = 0. \end{aligned} \tag{4.8}$$

Finally, at $s = (j+1)^-$, we have $[\, v_{(j+1)^-}(x_{j+1}) \,] = 1$, $j = 1, 2, \ldots, m$, and we obtain

$$\begin{aligned} &u_{j^+}\big(i\widetilde{\omega} M_{n,0} - G_{n,0}\big) + \sum_{r=1}^{n-1} u_{j+\frac{r}{n}}\big(i\widetilde{\omega} M_{n,r} - G_{n,r}\big) \\ &+ u_{(j+1)^-}\big(i\widetilde{\omega} M_{n,n} + G_{n,n} + \lambda - 1\big) - \lambda u_{(j+1)^+} = 0. \end{aligned} \tag{4.9}$$

Due to periodicity at end nodes, $m(n-1)$ equations are obtained from (4.7) at internal element nodes and (4.8) and (4.9) provide m additional equations each, leading to a total of $m(n+1)$ discrete equations, instead of mn equations for the CG scheme.

4.2. The Fourier analysis. As for the CG case, a dispersion analysis is now performed for the DG schemes. For symmetry reasons, only selected discrete equations are considered, (4.7) provides $n-1$ equations at interior nodes $j+\frac{r}{n}$, $r = 1, 2, \ldots, n-1$, and two equations are obtained from (4.8) and (4.9) at the typical nodes j^+ and $(j+1)^-$, respectively. The periodic solutions read as follows:

(i) $u_{j+\frac{r}{n}} = \widehat{u}_r\, e^{ikx_{j+\frac{r}{n}}}$ at the $n-1$ interior nodes $j+\frac{r}{n}$ of e_j, $r = 1, 2, \ldots, n-1$,

(ii) $u_{j^+} = \widehat{u}_0\, e^{ikx_j}$ and $u_{(j+1)^-} = \widehat{u}_n\, e^{ikx_{j+1}}$ at the boundary nodes of e_j,

and hence $n+1$ amplitudes are involved, instead of n for the CG scheme.

Substituting $u_{j+\frac{r}{n}}$, u_{j^+}, and $u_{(j+1)^-}$ in (4.7) leads to

$$\widehat{u}_0\, e^{-iqkh}\Big(i\widetilde{\omega}M_{q,0} - G_{q,0}\Big) + \sum_{r=1}^{n-1} \widehat{u}_r\, e^{i(r-q)kh}\,\big(i\widetilde{\omega}M_{q,r} - G_{q,r}\big)$$
$$+\,\widehat{u}_n\, e^{i(n-q)kh}\Big(i\widetilde{\omega}M_{q,n} - G_{q,n}\Big) = 0 \tag{4.10}$$

for $q = 1, 2, \ldots, n-1$, and substituting in (4.8) yields

$$\widehat{u}_0\,\Big(i\widetilde{\omega}M_{0,0} + G_{0,0} + \lambda\Big) + \sum_{r=1}^{n-1} \widehat{u}_r\, e^{irkh}\,\Big(i\widetilde{\omega}M_{0,r} - G_{0,r}\Big)$$
$$+\,\widehat{u}_n\,\Big(e^{inkh}\Big(i\widetilde{\omega}M_{0,n} - G_{0,n}\Big) + 1 - \lambda\Big) = 0. \tag{4.11}$$

Finally, substituting $u_{j+\frac{r}{n}}$, u_{j^+}, and $u_{(j+1)^-}$ in (4.9), we obtain

$$\widehat{u}_0\Big(e^{-inkh}\big(i\widetilde{\omega}M_{n,0} - G_{n,0}\big) - \lambda\Big) + \sum_{r=1}^{n-1} \widehat{u}_r\, e^{-i(n-r)kh}\,\Big(i\widetilde{\omega}M_{n,r} - G_{n,r}\Big)$$
$$+\,\widehat{u}_n\,\Big(i\widetilde{\omega}M_{n,n} + G_{n,n} + \lambda - 1\Big) = 0. \tag{4.12}$$

Equations (4.10)–(4.12) lead to an $(n+1)\times(n+1)$ matrix system

$$(i\widetilde{\omega}\mathcal{M} - \mathcal{G} - \widehat{\Lambda})\,\widehat{\mathbf{U}} = 0 \tag{4.13}$$

for amplitudes $\widehat{\mathbf{U}} = (\widehat{u}_0, \widehat{u}_1, \widehat{u}_2, \ldots, \widehat{u}_n)$, and $\widehat{\Lambda}$ is the $(n+1)\times(n+1)$ matrix whose first and last columns read as $\big(1-\lambda, 0, \ldots, 0, \lambda\, e^{inkh}\big)^T$ and $\big((\lambda-1)\, e^{-inkh}, 0, \ldots, 0, -\lambda\big)^T$, respectively, and all the remaining entries of $\widehat{\Lambda}$ are zero. A nontrivial solution $\widehat{\mathbf{U}}$ is obtained if $\det(i\widetilde{\omega}\mathcal{M} - \mathcal{G} - \widehat{\Lambda}) = 0$, leading to a polynomial in $\widetilde{\omega}$ of degree $n+1$, instead of n for the CG method. This discrepancy is due to an extra amplitude in the DG case, reflecting the discontinuous aspect of the method.

THEOREM 4.1. *For DG approximations of degree $n \le 10$, the dispersion relation obtained from* (4.13) *is the polynomial of degree $n+1$,*

$$P_{n+1}^{DG}(\widetilde{\omega}) = \widetilde{\omega}^{n+1} + \sum_{j=1}^{\lfloor\frac{n+1}{2}\rfloor} \tilde{b}_{n-2j+1}\,\widetilde{\omega}^{n-2j+1} + \sum_{j=0}^{\lfloor\frac{n}{2}\rfloor} \tilde{b}_{n-2j}\,\widetilde{\omega}^{n-2j} = 0, \tag{4.14}$$

with

$$\frac{\tilde{b}_{n-2j+1}}{a_{n-2j+1}} = \frac{(n+1)^2}{2jn} + (-1)^n\,\frac{(n+1)}{n}\Big(\cos(nkh) + i\,(2\lambda-1)\sin(nkh)\Big),$$
$$\frac{\tilde{b}_{n-2j}}{a_{n-2j}} = -i(2\lambda-1)\,\frac{(n+1)^2}{(2j+1)n} + (-1)^n\,\frac{(n+1)}{n}\Big(i(2\lambda-1)\cos(nkh) - \sin(nkh)\Big).$$

These results have been obtained using Maple. We conjecture they also hold for all $n > 10$.

The solutions of $P_{n+1}^{DG}(\widetilde{\omega})$ share a few properties obtained for $\widetilde{\omega}$ in section 3. For $\lambda = 1/2$, Theorem 3.1 holds with $\widehat{\mathcal{G}}$ and $\widehat{\mathcal{M}}$ replaced by $\mathcal{G} + \widehat{\Lambda}$ and $\mathcal{M}$, respectively. For all $\lambda \in [1/2, 1]$, Corollary 2 holds with odd and even n exchanged, and the group velocity $\partial\widetilde{\omega}/\partial kh$ is also $-(2n+1)$ at the end of the spectrum, as for the CG case. However, for $\lambda \in]1/2, 1]$, Theorem 3.1(i) and (iii) do not hold since $\mathcal{G} + \widehat{\Lambda}$ is no longer skew-Hermitian, and the solutions of $P_{n+1}^{DG}(\widetilde{\omega})$ are now complex numbers.

4.3. The presence of erratic modes. We let $\mathcal{G}_\lambda := \mathcal{G} + \Lambda$, where Λ is the $(n+1) \times (n+1)$ diagonal matrix whose diagonal entries starting in the upper left corner are $(1-\lambda, 0, 0, \ldots, 0, 0, -\lambda)$.

THEOREM 4.2. *The DG scheme does not admit any erratic stationary mode if and only if* $\lambda = 1$.

Proof. Letting $\omega = 0$ in (4.6) leads to computing the kernel of the discrete gradient operator. For symmetry reasons, only selected discrete equations are considered over interval e_j, $j = 1, 2, \ldots, m$, namely $n-1$ equations at interior nodes $j + \frac{q}{n}$, $q = 1, 2, \ldots, n-1$, and one equation at each typical boundary element node j^+ and $(j+1)^-$. From (4.7)–(4.9), a system of $n+1$ equations is then obtained with the $n+3$ unknowns $u_{j^-}, u_{j^+}, u_{j+\frac{1}{n}}, u_{j+\frac{2}{n}}, \ldots, u_{j+\frac{n-1}{n}}, u_{(j+1)^-}, u_{(j+1)^+}$:

$$\mathcal{G}_\lambda \left(u_{j^+}, u_{j+\frac{1}{n}}, \ldots, u_{j+\frac{n-1}{n}}, u_{(j+1)^-}\right)^T = \left((1-\lambda)u_{j^-}, 0, \ldots, 0, -\lambda u_{(j+1)^+}\right)^T, \tag{4.15}$$

where the vector in the RHS of (4.15) has $n+1$ components. From [11] we obtain

$$\det(\mathcal{G}_\lambda) = \det(\mathcal{G} + \Lambda) = \sum_{p=0}^{n+1} \sum_{q,r=0}^{n+1} (-1)^{s(q)+s(r)} \det(\mathcal{G}[q|r]) \det(\Lambda[q|r]), \tag{4.16}$$

where for a particular p, the inner sum is over strictly increasing integer sequences q and r of length p chosen from 1 to $n+1$; $\mathcal{G}[q|r]$ is the p–square submatrix of $\mathcal{G}$ lying in rows q and columns r; $\Lambda[q|r]$ is the $(n+1-p)-$square submatrix of Λ lying in rows complementary to q and columns complementary to r; and $s(q)$ is the sum of the integers in q. Further, $\det(\mathcal{G}_\lambda) = \det(\Lambda)$ for $p = 0$, and for $p = n+1$ it becomes $\det(\mathcal{G}_\lambda) = \det(\mathcal{G})$. The proof of (4.16) is a consequence of the linearity of the determinant in each row of the matrix, and the Laplace expansion theorem.

Theorem 3.3(i) yields $\det(\mathcal{G}) = 0$, and $\det(\Lambda) = 0$ for $n > 1$ and $\det(\Lambda) = \lambda(\lambda - 1)$ when $n = 1$ by definition of Λ. Further, because Λ has only two nonzero entries, the only summands that survive in the RHS of (4.16) correspond to $p = n$ and $p = n-1$:

(i) For $p = n$, two cases need to be considered. The first case is $q = r = \{2, n+1\}$, leading to $(1-\lambda)\det(\mathcal{G}^a)$, where $\mathcal{G}^a$ is an $n \times n$ matrix obtained from $\mathcal{G}$ by removing the first row and the first column. The second case is $q = r = \{1, n\}$, leading to $-\lambda \det(\mathcal{G}^b)$, where $\mathcal{G}^b$ is an $n \times n$ matrix obtained from $\mathcal{G}$ by removing the last row and the last column.

(ii) For $p = n-1$, only the case $q, r = \{2, n\}$ needs to be considered and it yields $\lambda(\lambda-1)\det \mathcal{G}^c$, where $\mathcal{G}^c$ is an $(n-1) \times (n-1)$ matrix obtained from $\mathcal{G}$ by removing the first and last rows and the first and last columns.

For $n > 1$, we then obtain $\det(\mathcal{G}_\lambda) = (1-\lambda)\det(\mathcal{G}^a) - \lambda \det(\mathcal{G}^b) + \lambda(\lambda-1)\det \mathcal{G}^c$. Manipulating rows and columns of matrices $\mathcal{G}^a, \mathcal{G}^b$, and $\mathcal{G}^c$ and employing properties (3.5)–(3.7) leads to $\det(\mathcal{G}^b) = (-1)^n \det(\mathcal{G}^a)$, $\det(\mathcal{G}^c) = \det(\mathcal{G}^a) - \det(\mathcal{G}^b)$, and $\det(\mathcal{G}^a) = \det(\mathcal{G}^\diamond)$, where $\mathcal{G}^\diamond$ is defined in (3.21). Finally, this yields

$$\det(\mathcal{G}_\lambda) = \left(1 - 2\lambda + \lambda^2(1 + (-1)^{n+1})\right) \det(\mathcal{G}^\diamond). \tag{4.17}$$

For $n = 1$, we obtain $\det(\mathcal{G}_\lambda) = \lambda^2 - \lambda + \frac{1}{2} \neq 0$ as $\lambda \in \mathbb{R}$. For $n > 1$, in (4.17) we deduce $\det(\mathcal{G}_\lambda) = (1 - 2\lambda)\det(\mathcal{G}^\diamond)$ for n even and $\det(\mathcal{G}_\lambda) = 2(\lambda^2 - \lambda + \frac{1}{2})\det(\mathcal{G}^\diamond)$ for n odd. From Theorem 3.3(ii), $\mathcal{G}^\diamond$ is full rank and $\mathcal{G}_1$ is hence invertible, further because the RHS of (4.15) only contains $u_{(j+1)^+}$ (and not u_{j^-}), the case $\lambda = 1$ does not allow the existence of a stationary erratic mode, and the kernel of $\mathcal{G}_1$ only contains the geostrophic mode of the form $(1, 1, 1, \ldots, 1)^T$. For $\lambda \neq 1$, as for the CG case, two vectors are solutions of (4.15): the geostrophic mode and the erratic stationary mode. The latter is computed from (4.15) for $\lambda = 1/2$ over interval e_j, $j = 1, 2, \ldots, m$, in Figure 13, and we obtain $(u_{j^+}, u_{(j+1)^-}) = (-1, 1)$ for $n = 1$, $(u_{j^+}, u_{j+\frac{1}{2}}, u_{(j+1)^-}) = (1, -\frac{1}{2}, 1)$ for $n = 2$, and $(u_{j^+}, u_{j+\frac{1}{3}}, u_{j+\frac{2}{3}}, u_{(j+1)^-}) = (-1, \frac{11}{27}, -\frac{11}{27}, 1)$ for $n = 3$. ∎

4.4. The presence of spectral gaps. Rearranging terms in (4.14) yields

$$\begin{aligned}\cos(nkh)\Big(R_{n-1}(\widetilde{\omega}) + i(2\lambda-1)Q_n(\widetilde{\omega})\Big) + \sin(nkh)\Big(i(2\lambda-1)R_{n-1}(\widetilde{\omega}) - Q_n(\widetilde{\omega})\Big)\\ + T_{n+1}(\widetilde{\omega}) + i(2\lambda-1)S_n(\widetilde{\omega}) = 0,\end{aligned} \tag{4.18}$$

where $T_{n+1}(\widetilde{\omega}) = (-1)^n \frac{n}{n+1}\widetilde{\omega}^{n+1} + (-1)^n(n+1)\sum_{j=1}^{\lfloor \frac{n+1}{2} \rfloor} \frac{a_{n-2j+1}}{2j}\widetilde{\omega}^{n-2j+1}$.

When $\lambda = 1/2$, the roots $\widetilde{\omega}_j$, $1 \leq j \leq n+1$, of $P^{DG}_{n+1}(\widetilde{\omega})$ in (4.18) do not exist if

$$|T_{n+1}(\widetilde{\omega})| > \sqrt{Q_n^2(\widetilde{\omega}) + R_{n-1}^2(\widetilde{\omega})}, \tag{4.19}$$

and $\widetilde{\omega}$ exhibits discontinuities or gaps, as in Corollary 3. Floor(n/2) pairs of gaps occur when $n \geq 2$, and Table 2 gives the intervals on which they exist for $n = 1, 2, \ldots, 7$.

TABLE 2

Intervals of values of $\widetilde{\omega}$, in the case $n = 1, 2, \ldots, 7$, for which the condition (4.19) is satisfied, namely when $\widetilde{\omega}$ cannot be computed. There is no gap when $n = 1$.

n	$\mathcal{I}_{\widetilde{\omega}}(n)$
2	[1.152 , 1.611]
3	[1.601 , 2.509]
4	[0.7005 , 0.7098] ∪ [1.877 , 3.217]
5	[1.1222 , 1.1722] ∪ [2.086 , 3.871]
6	[0.48575 , 0.48587] ∪ [1.399 , 1.513] ∪ [2.270 , 4.510]
7	[0.83858 , 0.84071] ∪ [1.597 , 1.788] ∪ [2.445 , 5.145]

When $\lambda = 1$, we have $\widetilde{\omega} \in \mathbb{C}$, and at least for $n \leq 10$ we do not observe gaps. In fact, we see graphically the disappearance of gaps when $\lambda \geq C > 1/2$, where C ranges from approximately 0.65 for $n = 2$ to approximately 0.75 for $n = 4$, for example.

4.5. The graphs of the dispersion relation. For $\lambda = 1/2$, the $n+1$ solutions of $P^{DG}_{n+1}(\widetilde{\omega})$ are shown in the first line of Figure 14 for $n = 1, 3, 5$ and of Figure 15 for $n = 2, 4, 6$. As in section 3, each solution is given its own color and it is valid over only a branch, and the union of all branches gives the solution $\widetilde{\omega}_S$, which can be constructed from $P^{CG}_{n+1}(\widetilde{\omega})$ using the branch selection procedure of section 3. Similar reconstructed u_h are obtained (not shown) based on eigenvectors. The solution $\widetilde{\omega}_S$ is displayed in the second line of Figures 14 to 15. As for the CG case, the spectral gaps computed in Table 2 are clearly visible in Figures 14 to 15. Note that for graphical purposes we now consider $kh \in [-\frac{n+1}{n}\pi, \frac{n+1}{n}\pi]$ since the indices of the local degrees of freedom need to be redistributed on a uniform and regular grid.

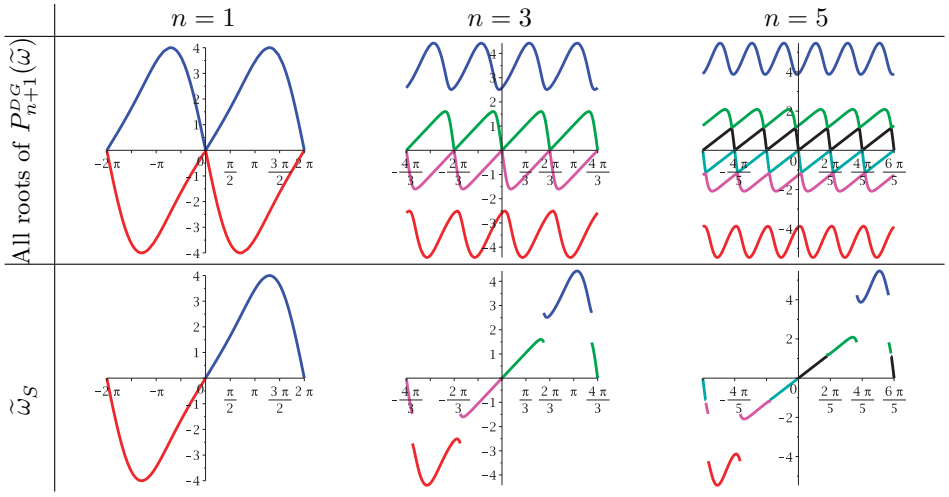

FIG. 14. *Phase $\widetilde{\omega}$ for the DG centered scheme ($\lambda = 1/2$) in the case $n = 1, 3, 5$. For a given n, the first line corresponds to all solutions of $P_{n+1}^{DG}(\widetilde{\omega})$ and $\widetilde{\omega}_S$ is given in the second line.*

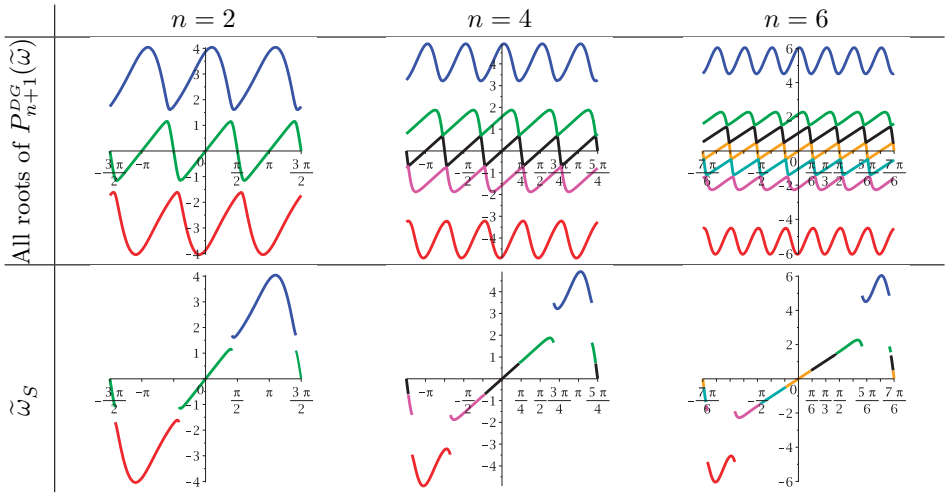

FIG. 15. *As for Figure 14 but in the case $n = 2, 4, 6$.*

In the case $\lambda = 1$, the real and imaginary parts of the solutions of $P_{n+1}^{DG}(\widetilde{\omega})$ are shown in the left column of Figure 16 for $n = 1, 3$ and of Figure 17 for $n = 2, 4$. The determination of branches is structurally quite different than found previously. The real part of the reconstructed solutions for $n = 3$ at $kh = \pi/24$ is shown in Figure 18. Contrary to the CG and centered DG cases, where the two solutions that cross at $kh = 0$ for odd n do not exchange roles for $kh = -\pi/24$, we obtain identical reconstructions when $\lambda = 1$. Since the reconstructed solution is unique only up to multiplication by a complex constant, this yields the observed shift in phase compared to Figure 8 for the CG case. The solution $\widetilde{\omega}_S$ is displayed in Figures 16 to 17 in the right column and in Figure 19 for $n = 5, 7, 10$. Unlike the CG and DG centered cases, there are no spectral gaps for any of these n, and the $(n+1)$ branches are connecting at $kh = \frac{p\pi}{n}$, $p = 1, 2, \ldots, n$. Further, as n increases, the damping becomes increasingly scale-

selective and localized to the high-frequency part of the spectrum, and also increases in strength. The nonzero imaginary part of the solution at the end of the spectrum now prevents the presence of an erratic stationary mode.

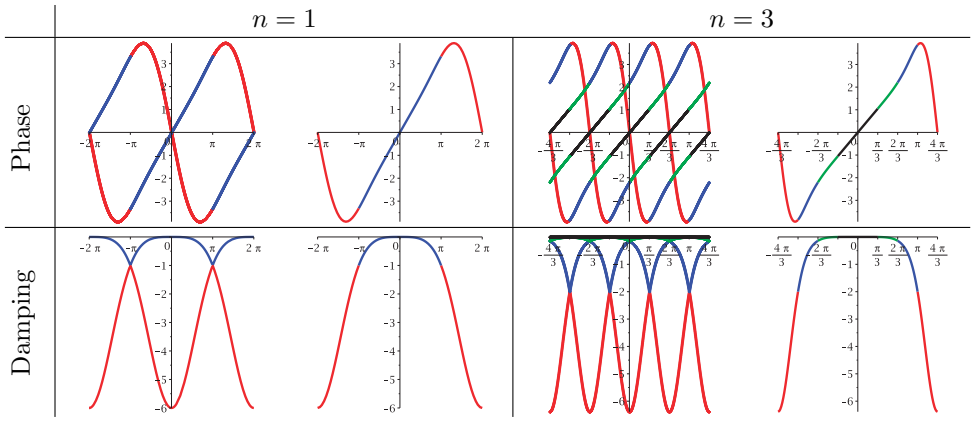

FIG. 16. *Phase* ($\Re(\widetilde{\omega})$) *and damping* ($-\Im(\widetilde{\omega})$) *of* $\widetilde{\omega}$ *for the DG upwind scheme* ($\lambda = 1$) *in the case* $n = 1$ *and* $n = 3$. *For a given* n, *the left column corresponds (for both phase and damping) to all solutions of* $P_{n+1}^{DG}(\widetilde{\omega})$ *and* $\widetilde{\omega}_S$ *is given in the right column.*

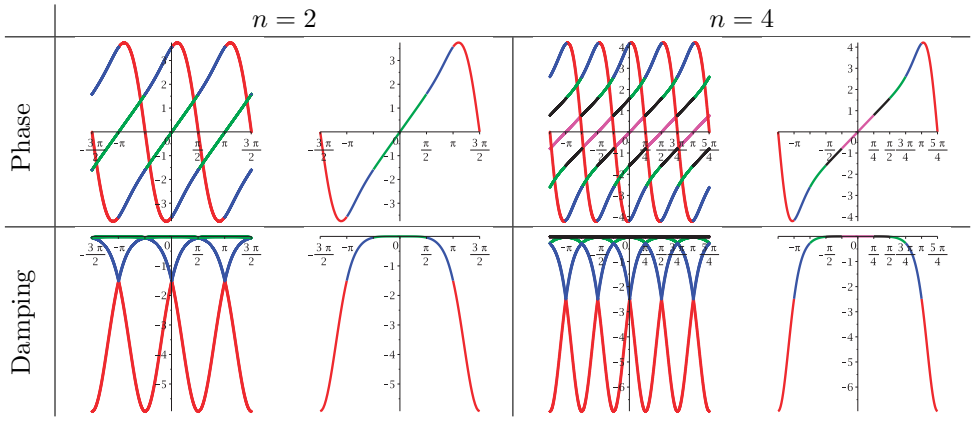

FIG. 17. *As for Figure* 16 *but in the case* $n = 2, 4$.

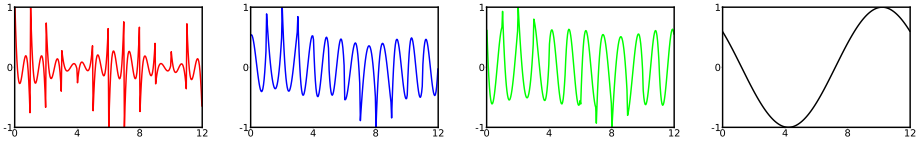

FIG. 18. *Reconstructed solution for DG upwind with* $n = 3$ *at* $kh = \pm\pi/24$.

5. Concluding remarks. The transport equation has been discretized for the CG and DG methods using high-degree interpolating polynomials, and the discrete equations were analyzed through Fourier analyses. Contrary to previous Fourier approaches, explicit analytical formulas have been obtained for the dispersion relations.

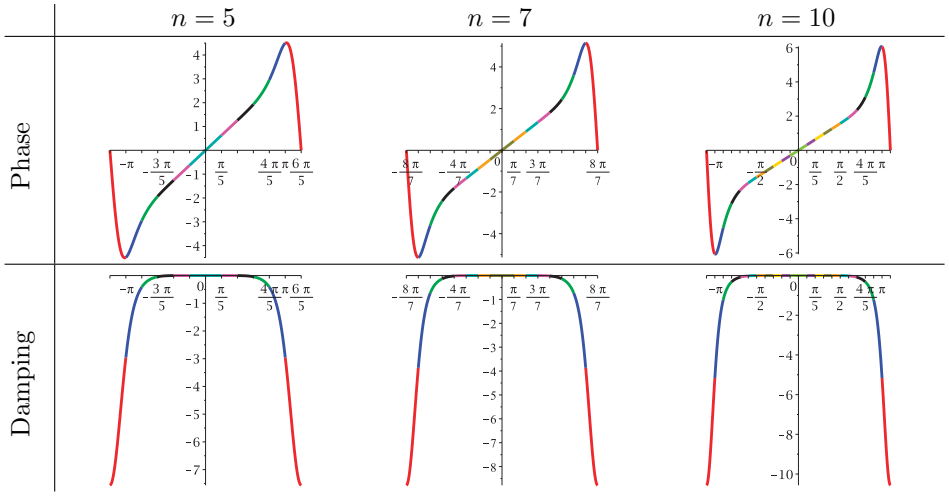

FIG. 19. *As for Figure* 16 *but for* $\widetilde{\omega}_S$ *in the case* $n = 5, 7, 10$.

These have permitted us, for the first time to our knowledge, to characterize analytically the presence of discontinuities or gaps in the dispersion relation for the CG and centered DG methods. We have also proposed a branch selection procedure to remove the mathematical artifacts generated by the Fourier method, leading to a single-valued dispersion relation. Finally, the existence of erratic stationary modes has been proven for the CG and centered DG methods. Conversely, upwind DG was shown to have neither spectral gaps nor an erratic stationary mode.

REFERENCES

[1] M. AINSWORTH, *Dispersive and dissipative behaviour of high order discontinuous Galerkin finite element methods*, J. Comput. Phys., 198 (2004), pp. 106–130.
[2] M. AINSWORTH, *Dispersive behaviour of high order finite element schemes for the one-way wave equation*, J. Comput. Phys., 259 (2014), pp. 1–10.
[3] C. ELDRED AND D. Y. LE ROUX, *Dispersion analysis of compatible Galerkin schemes for the* 1*d shallow water model*, J. Comput. Phys., 371 (2018), pp. 779–800.
[4] C. ELDRED AND D. Y. LE ROUX, *Dispersion analysis of compatible Galerkin schemes on quadrilaterals for shallow water model*, J. Comput. Phys., 387 (2019), pp. 539–568.
[5] A. ERN AND J. L. GUERMOND, *Theory and Practice of Finite Elements*, Springer, New York, 2004.
[6] J. S. HESTHAVEN AND T. WARBURTON, *Insight through theory*, in Nodal Discontinuous Galerkin Methods, Texts Appl. Math. 54, Springer, New York, 2008, pp. 75–113.
[7] F. Q. HU AND H. L. ATKINS, *Eigensolution analysis of the discontinuous Galerkin method with nonuniform grids*, J. Comput. Phys., 182 (2002), pp. 516–545.
[8] F. Q. HU, M. Y. HUSSAINI, AND P. RASETARINERA, *An analysis of the discontinuous Galerkin method for wave propagation problems*, J. Comput. Phys., 151 (1999), pp. 921–946.
[9] L. KRIVODONOVA AND R. QIN, *An analysis of the spectrum of the discontinuous Galerkin method*, Appl. Numer. Math., 64 (2013), pp. 1–18.
[10] P. H. LEBLOND AND L. A. MYSAK, *Waves in the Ocean*, Elsevier, Amsterdam, 1978.
[11] M. MARCUS, *Determinants of sums*, College Math. J., 21 (1990), pp. 130–135.
[12] T. MELVIN, A. STANIFORTH, AND J. THUBURN, *Dispersion analysis of the spectral element method*, Q. J. R. Meteorol. Soc., 138 (2012), pp. 1934–1947.
[13] G. MENGALDO, R. C. MOURA, B. GIRALDA, J. PEIRÓ, AND S. J. SHERVIN, *Spatial eigensolution analysis of discontinuous Galerkin schemes with practical insights for under-resolved computations and implicit LES*, Comput. & Fluids, 169 (2018), pp. 349–364.
[14] R. C. MOURA, M. AMAN, J. PEIRÓ, AND S. J. SHERVIN, *Spatial eigenanalysis of spectral/hp*

continuous Galerkin schemes and their stabilisation via DG-mimicking spectral vanishing viscosity: Application to high Reynolds number flows, J. Comput. Phys., 406 (2020), 109112.

[15] R. C. Moura, S. J. Shervin, and J. Peiró, *Linear dispersion-diffusion analysis and its application to under-resolved turbulence simulations using discontinuous Galerkin spectral/hp methods*, J. Comput. Phys., 298 (2015), pp. 695–710.

[16] A. H. Schatz, I. H. Sloan, and L. B. Wahlbin, *Superconvergence in finite element methods and meshes that are locally symmetric with respect to a point*, SIAM J. Numer. Anal., 33 (1996), pp. 505–521, https://doi.org/10.1137/0733027.

[17] J. C. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*, 2nd ed., SIAM, Philadelphia, 2004, https://doi.org/10.1137/1.9780898717938.

[18] P. A. Ullrich, D. R. Reynolds, J. E. Guerra, and M. A. Taylor, *Impact and importance of hyperdiffusion on the spectral element method: A linear dispersion analysis*, J. Comput. Phys., 375 (2018), pp. 427–446.