Universidad Autónoma de Madrid
Facultad de Ciencias
Departamento de Matemáticas

PROPIEDADES DE PROPAGACIÓN PARA ALGUNAS APROXIMACIONES POR MÉTODOS DE GALERKIN DISCONTINUOS Y ELEMENTOS FINITOS CLÁSICOS DE ORDEN SUPERIOR DE LA ECUACIÓN DE ONDAS

Memoria para optar al título de Doctor en Ciencias Matemáticas

Presentada por:
Aurora-Mihaela MARICA

Dirigida por:
Enrique ZUAZUA IRIONDO

MADRID
Septiembre 2010

Agradecimientos

Dedico esta memoria como un humilde regalo a mi hermana gemela, Neli-Mioara, cuya fuerza me ha movilizado en los momentos más difíciles de los últimos cinco años. También a mis padres y abuelos. A mi madre, por introducirme en el mundo fascinante de las matemáticas y por apoyarme a tomar decisiones radicales para mi futuro profesional. A mi padre, al menos por su eterno consejo "¡Lucha por ti!".

Quisiera manifestar mi gratitud hacia mi director de tesis, Enrique Zuazua, por haber aceptado la difícil tarea de dirigir a la estudiante más tímida de su carrera profesional, por su dedicación completa, a veces incluso fines de semana, por sus consejos oportunos y por la confianza que ha depositado en mí, cosas que han sido fundamentales para llevar a cabo esta memoria.

Quisiera darles las gracias a todos los miembros del tribunal de la tesis: Marius Tucsnak, Paola Antonietti, Blanca Ayuso, Julia Novo, Liviu Ignat, Carlos Castro y Daniel Faraco, por haber asumido la responsabilidad de evaluarla. Entre ellos, un agradecimiento especial para Liviu Ignat que ha sido el lector de la tesis doctoral. Sus recomendaciones han contribuido de manera considerable a conseguir una versión mejorada de la memoria. También a Blanca Ayuso, por la manera sincera de decir las cosas que muchas veces me ha hecho reflexionar sobre mi manera de actuar.

A mis profesores de la Universidad de Craiova y especialmente a Constantin Niculescu y Sorin Micu, por haber hecho posibles estos cinco años de investigación. En particular, quisiera agradecerle a Sorin su plena disponibilidad para trabajar conmigo durante las vacaciones.

A mis dos profesoras de matemáticas del colegio, Matilda Călugăru y Monalisa Găleteanu, excelentes pedagogas, cuyos años de enseñanza han contribuido de manera esencial a elegir el camino de las matemáticas.

A los miembros del Departamento de Matemáticas de la Universidad Autónoma de Madrid, donde pasé mis primeros tres años como estudiante de doctorado, y a los del Basque Center for Applied Mathematics de Bilbao, donde pasé los últimos dos, por haberme faciltado enormemente el trabajo y por haberme creado un ambiente favorable a la investigación.

A Cristian Cazacu, Răzvan Iagăr y Liviu Ignat, excelentes compañeros y amigos, con los cuales he compartido momentos agradables y cuyos consejos y apoyo me han ayudado a superar algunos momentos críticos.

Contents

Chapter 1

Introducción

A lo largo de esta memoria, analizamos las consecuencias de aumentar la complejidad de los esquemas numéricos sobre la aproximación de la propagación de ondas y los fenómenos de dispersión. Es bien sabido el hecho de que en general los métodos numéricos convergentes no sólo reproducen las soluciones de los modelos continuos como las Ecuaciones en Derivadas Parciales (EDPs), sino también pueden generar artefactos espúreos a altas frecuencias. Este fenómeno no constituye un impedimento para que la propiedad de convergencia en el sentido clásico del análisis numérico tenga lugar para condiciones iniciales y de contorno fijados, pero puede ser un inconveniente muy importante a la hora de utilizar clases más amplias de datos iniciales y soluciones, como pasa en el marco de los problemas de diseño óptimo y control. En estos contextos, se sabe que los requisitos clásicos de consistencia y estabilidad que garantizan la convergencia de las aproximaciones de las EDPs no bastan. Por ejemplo, en [59] se ha demostrado que los controles de la aproximación por diferencias finitas de la ecuación de ondas $1-d$ pueden diverger cuando el paso del mallado tiende a cero, incluso cuando la ecuación de ondas límite es controlable. De manera similar, en [31] se ha probado que las desigualdades de Strichartz clásicas no son uniformes con respeto al paso del mallado para las aproximaciones por diferencias finitas de la ecuación de Schrödinger. La primera contribución de esta tesis consiste en hacer un análisis sitemático de las patologías tipo altas frecuencias que son responsables de dichas inestabilidades.

Construimos de una manera rigurosa soluciones a altas frecuencias localizadas tanto en el espacio físico como en el de Fourier, de modo que sus soportes quedan localizados a lo largo de los rayos bicaracterísticos que se propagan con la llamada *velocidad de grupo* del esquema numérico, que puede degenerar cuando el paso del mallado converge a cero para los esquemas más típicos de aproximación de la ecuación de ondas. Esto està en contraste con el hecho de que todas las soluciones del modelo continuo correspondiente se propagan a una velocidad uniforme. Desarrollamos varias construcciones a altas frecuancias, que conducen todas a resultados similares y que están basados en análisis de ondas planas, *Gaussian beams* o expansiones WKB. Estas construcciones dan lugar a paquetes de ondas espúreos a altas frecuencias. Nuestro análisis hace más preciso el análisis de Fourier en N. Trefethen (cf. [54]) y los trabajos posteriores por L. Ignat, M. Negreanu y E. Zuazua (véase [32], [31], [45], entre otros). Desarrollamos estas construcciones primero en el contexto de las diferencias finitas, para luego considerar métodos numéricos más sofisticados basados en aproximaciones de Galerkin discontinuas (GD) y elementos finitos de orden superior. Como veremos, incluso al incrementar el orden de aproximación de los esquemas numéricos o relajándolos mediante un enfoque GD, las patologías a altas frecuancias persisten. Estos métodos son más complejos, en el sentido que involucran varias relaciones de dispersión. Cada una de estas relaciones de dispersión, una física y al menos una espúrea, tiene sus propios puntos críticos. Teniendo esto en cuenta y usando un desacoplamiento en el espacio de Fourier, se pueden construir paquetes de ondas a altas frecuencias localizados entorno a cada punto crítico del espectro, negando cualquier tipo de propiedad de propagación o dispersión uniforme. Probamos ciertas propiedades finas de estos paquetes de ondas en lo que respecta a su dependencia del

paso del mallado, como por ejemplo, la tasa de divergencia de la correspondiente constante de observabilidad o su falta de simetría. Para los métodos GD o las aproximaciones por elementos finitos clásicos de orden superior, existen incluso paquetes de ondas que se propagan en el sentido contrario al caso continuo, localizados en el diagrama espúreo.

Teniendo en cuenta estas patologías, en la literatura consagrada a desarrollar algoritmos numéricos estables para el control de ondas o a resolver ecuaciones dispersivas, se han diseñado varios mecanismos para reducir el efecto de las componentes en altas frecuencias, considerando, por ejemplo, subclases de datos iniciales filtrados usando un truncamiento de las altas frecuencias [58] en el espacio de Fourier, elementos finitos mixtos [17], algoritmos bimalla [32], [45], viscosidad numérica [52] o regularización de Tychonoff [26]. En esta tesis, adaptamos estos mecanismos de filtrado al caso de las aproximaciones DG o por elementos finitos clásicos de orden superior. Debido al comportamiento más complejo de estos esquemas, se requiere una atenta combinación de varias técnicas de filtrado para eliminar de una manera eficaz los efectos espúreos.

En esta memoria, presentamos resultados sobre cuatro temas:

1. **Ondas concentradas para la aproximación por diferencias finitas de la ecuación de ondas.** Desarrollamos una construcción rigurosa de paquetes de ondas en altas frecuencias para la semi-discretización por diferencias finitas de la ecuación de ondas. Usamos técnicas distintas (métodos asintóticos - WKB o Gaussian beams - y construcciones explícitas de datos iniciales exponencialmente concentrados para la ecuación de ondas semi-discreta) y mostramos la falta de uniformidad con respeto al paso del mallado en el modelo semi-discreto y, en particular, la divergencia a un orden polinomial arbitrario de la constante de observabilidad.

2. **Propiedades de propagación de las aproximaciones Galerkin discontinuas de la ecuación de ondas** $1-d$**.** Hacemos un estudio sistemático de las propiedades de propagación de las semi-discretizaciones por métodos GD de la ecuación de ondas $1-d$. En concreto, hacemos un análisis de Fourier cuidadoso de algunas clases de métodos GD en el caso $1-d$ y construimos soluciones numéricas concentradas en altas frecuencias, mostrando la falta de uniformidad en las propiedades de observabilidad para el modelo discreto. También diseñamos mecanismos de filtrado para la ecuación de ondas semi-discreta, basados en el truncamiento en Fourier y el algoritmo bimalla.

3. **Propiedades de propagación para las aproximaciones por funciones spline de la ecuación de ondas** $1-d$**.** Desarrollamos un análisis sistemático de las propiedades de propagación de los llamados k-refinamientos o métodos C^k, que consisten en usar funciones de base de tipo *spline* para aproximar las soluciones de la ecuación de ondas. Probamos que para cualquier grado de regularidad k de estos métodos, la velocidad de propagación de los paquetes de ondas sigue sin tener una cota inferior estrictamente positiva. Sin embargo, a medida que k aumenta, la relación de dispersión discreta se adapta mejor a la continua para un rango más extendido de números de onda.

4. **Propiedades de propagación para la semi-discretización por elementos finitos clásicos cuadráticos de la ecuación de ondas.** Realizamos un análisis de Fourier riguroso de la aproximación por elementos finitos clásicos cuadráticos de la ecuación de ondas $1-d$. Aparecen fenómenos patológicos similares a la semi-discretización por métodos GD. Diseñamos los correspondientes mecanismos de filtrado.

Al final de la tesis, incluimos una lista de problemas abiertos relacionados con los temas que tratamos a lo largo de la tesis que pueden constituir interesantes líneas de investigación en el futuro.

A continuación, describimos brevemente los aspectos más relevantes de los problemas estudiados, los resultados obtenidos y la metodología empleada.

CAPÍTULO 1. Aproximaciónes por diferencias finitas de la ecuación de ondas. El problema de construir paquetes de ondas que se propagan a la velocidad de grupo para las aproximaciones numéricas de las ecuaciones de ondas o transporte ha sido ya investigado, empezando con el artículo pionero de Trefethen [54]. Sin embargo, nuestro análisis es más profundo y da lugar a resultados más precisos. Nuestras construcciones están motivadas por temas de control, pero poseen también un interés intrínseco.

Dado $T > 0$ y el tamaño de la malla $h > 0$, consideramos el problema de Cauchy asociado a la ecuación de ondas semi-discretizada por diferencias finitas en un mallado uniforme de paso h

$$\begin{cases} \partial_t^2 u_{\mathbf{j}}(t) - \Delta_h u_{\mathbf{j}}(t) = 0, & \mathbf{j} \in \mathbb{Z}^d,\ t \in (0,T) \\ u_{\mathbf{j}}(0) = u_{\mathbf{j}}^0,\ \partial_t u_{\mathbf{j}}(0) = u_{\mathbf{j}}^1, & \mathbf{j} \in \mathbb{Z}^d, \end{cases} \tag{1.1}$$

con $\Delta_h u_{\mathbf{j}} = h^{-2} \sum_{k=1}^{d} \left(u_{\mathbf{j}+\mathbf{e}_k} - 2u_{\mathbf{j}} + u_{\mathbf{j}-\mathbf{e}_k} \right)$.

El problema (1.1) constituye una aproximación del problema de Cauchy asociado a la ecuación de ondas continua:

$$\begin{cases} \partial_t^2 u(x,t) - \Delta u(x,t) = 0, & x \in \mathbb{R}^d,\ t \in (0,T) \\ u(x,0) = u^0(x), \quad \partial_t u(x,0) = u^1(x), & x \in \mathbb{R}^d, \end{cases} \tag{1.2}$$

en una malla espacial uniforme de paso h. Evidentemente, $u_{\mathbf{j}}(t)$ denota una aproximación de $u(x,t)$ en el punto $x = x_{\mathbf{j}} = \mathbf{j}h$, $\mathbf{j} \in \mathbb{Z}^d$.

Introducimos también el gradiente discreto $\nabla_h = (\partial_{h,k})_{k=1,\dots,d}$, donde

$$\partial_{h,k} \overrightarrow{f} = \frac{\overrightarrow{f}_{\cdot+\mathbf{e}_k} - \overrightarrow{f}}{h},$$

$\overrightarrow{f} = (f_{\mathbf{j}})_{\mathbf{j}\in\mathbb{Z}^d}$ es una succesión cualquiera y $(\mathbf{e}_l)_{l=1}^d$ es la base canónica en $\mathbb{R}^d$. Para un conjunto abierto $O \subset \mathbb{R}^d$, definimos los espacios de Sobolev discretos

$$\ell^2(O) = \{ \overrightarrow{f} \text{ t.q. } \| \overrightarrow{f} \|_{\ell^2(O)}^2 := h^d \sum_{x_{\mathbf{j}} \in O} |f_{\mathbf{j}}|^2 < \infty \}$$

y

$$\dot{h}^1(O) = \{ \overrightarrow{f} \text{ t.q. } \| \overrightarrow{f} \|_{\dot{h}^1(O)}^2 := \sum_{k=1}^{d} \| \partial_{h,k} \overrightarrow{f} \|_{\ell^2(O)}^2 < \infty \}.$$

Si $\Gamma \subset O$ es una hipersuperficie, definimos el siguiente espacio de Sobolev discreto sobre Γ:

$$\ell^2(\Gamma) = \{ \overrightarrow{f} = (f_{\mathbf{j}})_{x_{\mathbf{j}} \in \Gamma} : \| \overrightarrow{f} \|_{\ell^2(\gamma)}^2 := h^{d-1} \sum_{x_{\mathbf{j}} \in \Gamma} |f_{\mathbf{j}}|^2 < \infty \}.$$

Sea $\Omega := \mathbb{R}^d \setminus B^d(0,1)$, donde $B^d(0,1)$ es la bola unidad d-dimensional. Nuestra motivación original es el comportamiento cuando $h \to 0$ de la constante

$$C_h(T) = \sup_{(\overrightarrow{u}^0, \overrightarrow{u}^1) \in \dot{h}^1(\mathbb{R}^d) \times \ell^2(\mathbb{R}^d)} \frac{\| \overrightarrow{u}^0 \|_{\dot{h}^1(\mathbb{R}^d)}^2 + \| \overrightarrow{u}^1 \|_{\ell^2(\mathbb{R}^d)}^2}{\int_0^T (\| \nabla_h \overrightarrow{u}(t) \|_{\ell^2(\Omega)}^2 + \| \partial_t \overrightarrow{u}(t) \|_{\ell^2(\Omega)}^2)\, dt}. \tag{1.3}$$

Para cualquier $T > 0$ y $h > 0$, la constante $C_h(T)$, llamada *constante de observabilidad*, da información de si las mediciones hechas en el dominio exterior Ω durante el intervalo temporal $(0,T)$ pueden ser eficientes al estimar la energía total de las soluciones de (1.1), y se puede probar que es finita. Sin embargo, es bien sabido que $C_h(T)$ diverge cuando $h \to 0$ para cualquier $T > 0$ finito (cf. [59]). Nuestro objetivo es analizar la tasa de divergencia. Se sabe que en el caso

continuo, para $T > 2$, la constante de observabilidad es finita ([7], [36]). Sin embargo, para las aproximaciones numéricas, la constante de observabilidad $C_h(T)$ explota a medida que $h \to 0$, para todo $T > 0$, lo que muestra la inestabilidad del esquema numérico en los problemas de control.

El problema de observabilidad que estudiamos a lo largo de este capítulo es interesante en sí mismo, y además tiene algunas aplicaciones en la teoría del control, en problemas inversos y de estabilización. En particular, por el Método de Unicidad de Hilbert (HUM), el problema (1.1) es equivalente a un problema de controlabilidad exacta a través de un control ubicado en el conjunto $\Omega = \mathbb{R}^d \setminus B^d(0,1)$. De manera más precisa, dicho problema de control se puede formular de la siguiente manera. Para todo $(\overrightarrow{u}^0, \overrightarrow{u}^1) \in \dot{\hbar}^1(h\mathbb{Z}^d) \times \ell^2(h\mathbb{Z}^d)$ y todo tiempo finito $T > 0$, hay que encontrar un control $(f_{\mathbf{j}}\chi_{|x_{\mathbf{j}}|>1}(\mathbf{j}))_{\mathbf{j}\in\mathbb{Z}^d}$ tal que la solución del problema no homogéneo

$$\begin{cases} \partial_t^2 u_{\mathbf{j}}(t) - \Delta_h u_{\mathbf{j}}(t) = f_{\mathbf{j}}\chi_{|x_{\mathbf{j}}|>1}(\mathbf{j}), & \mathbf{j} \in \mathbb{Z}^d, t \in (0,T) \\ u_{\mathbf{j}}(0) = u_{\mathbf{j}}^0,\ \partial_t u_{\mathbf{j}}(0) = u_{\mathbf{j}}^1, & \mathbf{j} \in \mathbb{Z}^d, \end{cases} \tag{1.4}$$

satisfaga $u_{\mathbf{j}}(T) = \partial_t u_{\mathbf{j}}(T) = 0$. Del hecho de que la constante de observabilidad $C_h(T)$ diverje de manera polinomial resulta que existen clases de datos iniciales en (1.4) para los cuales el control minimal f en $L^2(0,T,\ell^2(\Omega))$ diverge a una velocidad polinomial de orden arbitrario en $L^2(0,T;\ell^2(\Omega))$ cuando $h \to 0$.

Se pueden analizar problemas de control similares en dominios acotados, con controles ubicados en el interior o en la frontera del dominio. En concreto, para $h = 1/(N+1)$ y $\Gamma_0 \subset \Gamma = \partial\mathcal{I}$, $\mathcal{I} := (0,1)^d$, consideramos el problema semi-discreto de valores iniciales con control ubicado en la frontera

$$\begin{cases} \partial_t^2 u_{\mathbf{j}}(t) - \Delta_h u_{\mathbf{j}}(t) = 0, & x_{\mathbf{j}} \in \mathcal{I},\ t \in (0,T) \\ u_{\mathbf{j}}(t) = v_{\mathbf{j}}^h(t)\chi_{\Gamma_0}(\mathbf{j}), & x_{\mathbf{j}} \in \Gamma,\ t \in (0,T) \\ u_{\mathbf{j}}(0) = u_{\mathbf{j}}^0,\ \partial_t u_{\mathbf{j}}(0) = u_{\mathbf{j}}^1, & x_{\mathbf{j}} \in \mathcal{I}. \end{cases} \tag{1.5}$$

El problema de controlabilidad exacta requiere encontrar un control $\overrightarrow{v}^h \in L^2(0,T;\ell^2(\Gamma_0))$ tal que, para todo $(\overrightarrow{u}^0, \overrightarrow{u}^1) \in \ell^2(\mathcal{I}) \times \hbar^{-1}(\mathcal{I})$, la solución de (1.5) satisfaga $u_{\mathbf{j}}(T) = \partial_t u_{\mathbf{j}}(T) = 0$, para todo $x_{\mathbf{j}} \in \mathcal{I}$. Por $\hbar^{-1}(\mathcal{I})$ denotamos el dual de $\hbar_0^1(\mathcal{I})$, el análogo discreto de $H_0^1(\mathcal{I})$, definido por

$$\hbar_0^1(\mathcal{I}) = \{\overrightarrow{f} = (f_{\mathbf{j}})_{x_{\mathbf{j}}\in\mathcal{I}} : \|\overrightarrow{f}\|_{\hbar_0^1(\mathcal{I})} := \|\overrightarrow{f}\|_{\dot{\hbar}^1(\mathcal{I})} < \infty, f_{\mathbf{j}} = 0, x_{\mathbf{j}} \in \Gamma\}.$$

Por el Método de Unicidad de Hilbert (HUM), el problema de controlabilidad exacta (1.5) es equivalente a un problema de observabilidad para el problema semi-discreto adjunto,

$$\begin{cases} \partial_t^2 \phi_{\mathbf{j}}(t) - \Delta_h \phi_{\mathbf{j}}(t) = 0, & x_{\mathbf{j}} \in \mathcal{I},\ t \in (0,T) \\ \phi_{\mathbf{j}}(t) = 0, & x_{\mathbf{j}} \in \Gamma,\ t \in (0,T) \\ \phi_{\mathbf{j}}(0) = \phi_{\mathbf{j}}^0,\ \partial_t \phi_{\mathbf{j}}(0) = \phi_{\mathbf{j}}^1, & x_{\mathbf{j}} \in \mathcal{I}. \end{cases} \tag{1.6}$$

Más precisamente, la propiedad de controlabilidad exacta ocurre si y sólo si la siguiente constante de observabilidad es finita:

$$C_h(T) = \sup_{(\overrightarrow{\phi}^0, \overrightarrow{\phi}^1)\in\hbar_0^1(\mathcal{I})\times\ell^2(\mathcal{I})} \frac{\|\overrightarrow{\phi}^0\|_{\hbar_0^1(\mathcal{I})}^2 + \|\overrightarrow{\phi}^1\|_{\ell^2(\mathcal{I})}^2}{\sum\limits_{\mathbf{j}\in\Gamma_{h,0}} \int\limits_0^T \|\partial_{h,\nu}\overrightarrow{\phi}(t)\|_{\ell^2(\Gamma_0)}^2\, dt}, \tag{1.7}$$

donde el supremo se toma sobre el conjunto de todas las soluciones $\overrightarrow{\phi}$ de energía finita de (1.6), ν es el vector normal unitario exterior a Γ_0, y $\partial_{h,\nu}$ es la derivada normal discreta definida por

$$\partial_{h,\nu} f_{\mathbf{j}} = -\frac{1}{h} f_{\mathbf{j}-\nu}, \forall x_{\mathbf{j}} \in \Gamma_0,\ \overrightarrow{f} \in \hbar_0^1(\mathcal{I}).$$

En [33] y [58] se ha mostrado que si $d = 1$ o $d = 2$, $C_h(T) \to \infty$ cuando $h \to 0$ porque el *gap* de los autovalores localizados en un pequeño entorno de los números de onda $\{0, \pi/h\}^d \setminus \{0\}$

converge a cero cuando $h \to 0$. Además, se ha mostrado que un truncamiento en el espacio de Fourier es eficiente para recuperar la propiedad de observabilidad uniforme en todas las subclases de soluciones que involucran bajas frecuencias. En [45] y [32] se ha analizado la eficacia del método bimalla para las versiones $1-d$ y $2-d$ de (1.6) para reestablecer la uniformidad de la constante $C_h(T)$ cuando $h \to 0$. En [42] se lleva al cabo un análisis fino de los controles en (1.5) en el caso $1-d$. A través de un análisis fino de sucesiones biortogonales, Micu prueba la existencia de un control de orden $\exp(\sqrt{N})$. Por HUM, esto demuestra que $C_h(T)$ definido en (1.7) explota al menos de manera exponencial. Usando este resultado de [42], en la Proposition 1.2.1, pp. 7, mostramos la existencia de paquetes de ondas para los cuales la constante de observabilidad $C_h(T)$ en (1.3) explota de manera exponencial.

Una de nuestras contribuciones originales es encontrar una cota inferior para la tasa de divergencia de $C_h(T)$ que sea válida en cualquier dimension espacial para el problema de Cauchy (1.1) en todo el espacio euclídeo. Más concretamente, demostramos que $C_h(T)$ diverge de una manera polinomial de orden arbitrario, desarrollando varios construcciones de paquetes de ondas en altas frecuencias ligeramente distintos. Nuestros resultados generalizan los de [42], que muestran la divergencia exponencial de la constante de observabilidad para el problema (1.6) en $d = 1$, aunque nosotros no obtenemos una tasa de divergencia exponencial. Las construcciones a altas frecuencias que desarrollamos son las siguientes:

- Soluciones exactas explícitas generadas por datos iniciales regulares en la classe de Schwartz de soporte de orden h^α en el espacio físico, para algún $\alpha < 1$
- Expansiones WKB discretas
- Construcciones *Gaussian beams* discretas, que extienden al marco discreto la construcción de Ralston en [46].

En vista de su localización espacio-temporal, las soluciones que construimos para el problema de Cauchy (1.1) pueden ser fácilmente adaptadas para probar divergencia de orden polinomial arbitrario de las constantes de observabilidad para problemas de ondas discretos en dominios acotados con distintas condiciones de contorno.

También presentamos varios graficos de estos paquetes de ondas en altas frecuancias y hacemos un análisis cuidadoso de algunas propiedades cualitativas, especialmente las dispersivas. De esta manera, observamos y mostramos que, aunque los datos iniciales son simétricos, por ejemplo muestras de funciones gaussianas, la solución numérica corespondiente a tiempos próximos no preserva esta simetría. Esto se debe al hecho de que los paquetes de ondas que construimos viajan a lo largo de las características, pero no todos los frentes de ondas viajan a la misma velocidad. Esto hace que se rompa la simetría que hay en la solución continua correspondiente al mismo dato inicial. Otra propiedad que queremos enfatizar es la similitud con una función gaussiana. Mostramos que para $\alpha < 2/3$ y tiempos finitos, los paquetes de ondas que construimos a partir de datos iniciales gaussianos se pueden aproximar por un perfil gaussiano. Las simulaciones numéricas demuestran que para valores más grandes de α, esta semejanza con un perfil gaussiano se pierde, pero hasta ahora no tenemos una prueba rigurosa de esta propiedad.

Las contrucciones presentadas en este capítulo serán adaptadas a lo largo de los siguientes capítulos a esquemas numéricos más complejos, en particular los generados por métodos GD.

CAPÍTULO 2. Semi-discretizaciones por métodos de Galerkin discontinuos de la ecuación de ondas $1-d$. Los métodos Galerkin discontinuos (GD) son aproximaciones no conformes de las EDPs que producen soluciones que presentan discontinuidades a lo largo de las interfaces. Para asegurar la estabilidad de este tipo de aproximaciones, se suelen penalizar los saltos añadiendo flujos numéricos que involucran algunos parámetros de estabilización.

En el segundo capítulo de esta memoria, analizamos las propiedades de propagación de algunas semi-discretizaciones GD de la ecuación de ondas $1-d$. Se considera el caso más sencillo de las aproximaciones por polinomios de orden uno en una malla uniforme de tamaño h. Una semi-

discretización de este tipo se puede escribir en la forma de un sistema infinito de ODEs

$$M_h \overrightarrow{U}_{tt}(t) + R_h^s \overrightarrow{U}(t) = 0,$$

donde $\overrightarrow{U}(t) = (A_k(t), J_k(t))_{k\in\mathbb{Z}}$ es el vector de incógnitas constituido por dos tipos de componentes, $A_k(t)$ y $J_k(t)$, que representan la media y el salto de la solución numérica en el nodo x_k. Denotamos por R_h^s y M_h las matrices de rigidez y masa que son tridiagonales por bloques y dependen de h y del parámetro de estabilización $s > 1$.

Tomando transformadas de Fourier semi-discretas (TFSDs) en la variable espacial, el sistema de ODEs se puede transformar en un sistema de ODEs de tamaño 2×2 dependiente del parámetro de Fourier $\xi \in \Pi_h$:

$$\widehat{U}_{tt}^h(\xi,t) + S_h^s(\xi)\widehat{U}^h(\xi,t) = 0, \tag{1.8}$$

donde $S_h^s(\xi) = (M_h(\xi))^{-1}R_h^s(\xi)$ y $M_h(\xi), R_h^s(\xi)$ son matrices 2×2 que representan los símbolos de Fourier de las matrices M_h y R_h^s y dependen de la variable Fourier $\xi \in \Pi_h$. Las dos componentes del vector de incógnitas $\widehat{U}^h(\xi,t)$ representan las partes continua y de salto de la solución numérica, respectivamente. Evidentemente, para obtener la unicidad de la solución numérica, hay que complementar el sistema (1.8) con dos datos iniciales vectoriales $\widehat{U}^h(\xi,0)$ y $\widehat{U}_t^h(\xi,0)$.

Para el métodos GD llamado *symmetric interior penalty* (SIPG) introducido en [5], el análisis de Fourier que desarrollamos demuestra la coexistencia de dos diagramas de dispersión, que son de hecho las raíces cuadradas de los dos autovalores de la matriz $S_h^s(\xi)$: una física (acústica), que es la más pequeña de los dos diagramas y cuyo comportamento es similar a la relación de dispersión para elementos clásicos de orden uno, y una espúrea (óptica), que tiende a infinito cuando $h \to 0$ o el parámetro de estabilización tiende a infinito.

Aunque el diagrama de dispersión física para los métodos SIPG está mucho más cerca de la continua que cuando se usan aproximaciones estándar por diferencias finitas o elementos clásicos lineales, el análisis de la velocidad de grupo muestra que ambas ramas, acústica y óptica, presentan puntos singulares: uno ubicado en el diagrama de dispersión física en el número de onda π/h y al menos otros dos situados en el diagrama de dispersión espúrea, en los números de onda 0 y π/h. Tener puntos críticos en los diagramas de dispersión en el número de onda más alto π/h es un fenómeno clásico, bien conocido para los esquemas en diferencias finitas y elementos finitos lineales. La singularidad en el número de onda 0 en la rama óptica es menos habitual. Esto se debe a las propiedades de simetría del diagrama espúreo con respeto a $\xi = 0$, ya que la relación de dispersión es una expresión racional de polinomios trigonométricos, y al hecho de que el símbolo de Fourier de la matriz de masa $M_h(\xi)$ sea no singular.

Para cada uno de estos puntos singulares, podemos adaptar la construcción de soluciones concentradas del primer capítulo. Para construir paquetes de ondas que se propagan con la velocidad de grupo corespondiente sólo a una de las relaciones de dispersión e impedir la interacción entre los dos modos, se usa un desacoplamiento en el espacio de Fourier. Esto produce soluciones que involucran sólo una relación de dispersión. Además, hay que tener en cuenta el hecho de que, en el diagrama de dispersión espúrea, en general, los paquetes de ondas viajan en la dirección no física, lo que los hace doblemente no físicos: por la velocidad de propagación y también por la dirección de propagación.

Otro de los objetivos alcanzados es el diseñar mecanismos de filtrado apropiados para esta clase de semi-discretizaciones más complejas, dirigidos a recuperar la uniformidad de la constante de observabilidad con respeto al paso del mallado h en subespacios correspondientes de soluciones numéricas.

Se analizan dos estrategias principales:

A) Usando un desacoplamiento en el espacio de Fourier y una elección apropiada de los datos iniciales, se trabaja con soluciones de la semi-discretizaciones SIPG que involucran sólo la relación de dispersión física, que tiene únicamente un punto singular. Por lo tanto, se pueden aplicar los algoritmos de filtrado clásicos para los datos iniciales: *a) un truncamiento en el espacio Fourier de*

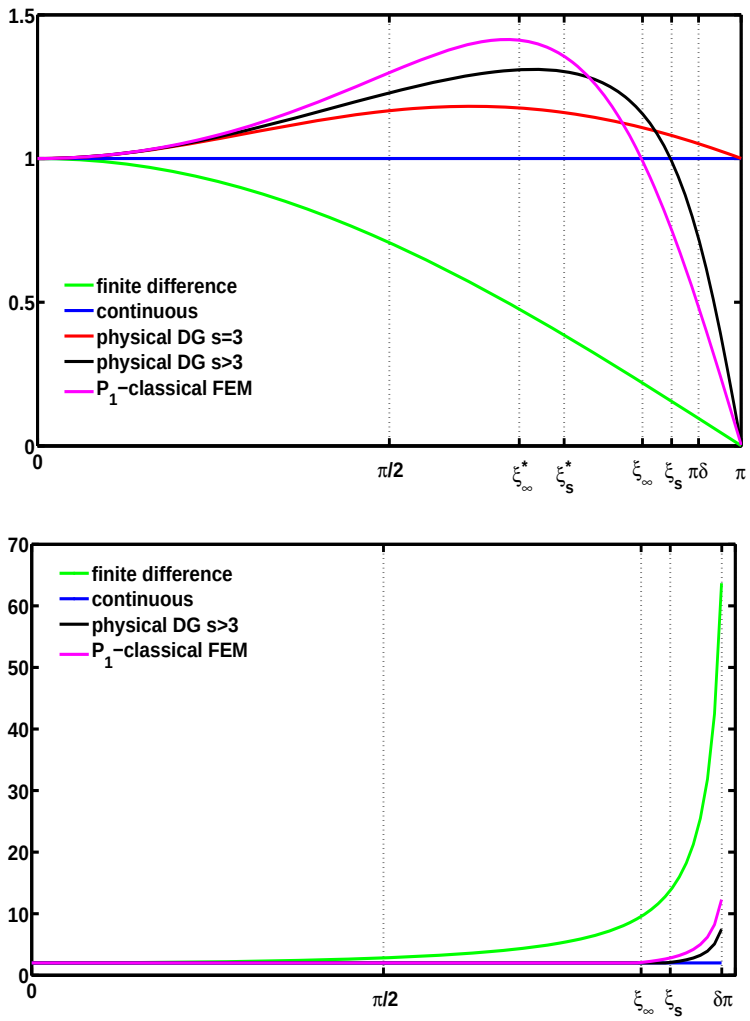

Figure 1.1: Velocidad de grupo versus tiempo de observabilidad óptimo correspondientes a las semi-discretizaciones de la ecuación de ondas $1-d$ por diferencias finitas, el modo físico en el método SIPG y elementos finitos clásicos lineales.

las altas frecuencias: considerar datos iniciales para los cuales el soporte de las correspondientes TFSDs no contiene los puntos críticos de las correspondientes relaciones de dispersión, por ejemplo, $\xi \in [-\pi\delta/h, \pi\delta/h]$, con $\delta \in (0,1)$ o *b) un algoritmo bimalla*: considerar datos iniciales obtenidos por interpolación lineal a partir de una malla más gruesa.

B) La segunda estrategia está diseñada para que, dados los datos iniciales del modelo continuo, las correpondientes proyecciones sobre la malla numérica se construyen de tal manera que la mayor parte de su energía se concentre en la rama física de la diagrama de dispersión. El caso más sencillo de una proyección de este tipo es cuando $J_k(0) = \partial_t J_k(0) = 0$, para todo $k \in \mathbb{Z}$, es decir, cuando los datos iniciales numéricos son tal que sus saltos se anulen. Esto produce un peso para las componentes espúreas en la expresión de la TFDS de la solución que se anula en el número de onda $\xi = 0$, y elimina de esta manera el comportamiento patológico en dicho número de onda. Para eliminar el comportamiento patológico en $\xi = \pi/h$, hay que aplicar o bien *a) un truncamiento en Fourier de las altas frecuencias* o *b) un algoritmo bimalla* sobre los datos iniciales

físicos $(A_k(0), \partial_t A_k(0))_{k\in\mathbb{Z}}$.

Desde un punto de vista práctico, la estrategia más factible y útil consiste en primero elegir los saltos nulos en los datos iniciales numéricos, y luego aplicar el filtrado bimalla. Esto asegura que la energía de las soluciones se concentra en las bajas frecuencias de la rama física. Además, tiene la ventaja de ser aplicable en la malla física y no requiere la TFSD.

El análisis desarrollado en este capítulo muestra que el comportamiento global de las soluciones numéricas obtenidas por los métodos GD es más complejo que el correspondiente a la mayoría de los métodos numéricos clásicos como diferencias finitas o elementos clásicos lineales. Ahora bien, después de filtrar apropiadamente los datos iniciales considerados, la solución obtenida resulta estar más cerca de la de la ecuación de ondas original que las que producen los métodos más clásicos. El tiempo de observabilidad dado por la Condición de Control Geométrico (que establece que el tiempo de observabilidad debe ser el necesario para que todos los rayos bi-característicos entren en la región de observación) es $2/v$, donde v es la velocidad de propagación mínima de las componentes de Fourier involucradas en el paquete de ondas correspondiente. Para los esquemas más clásicos, como, por ejemplo, diferencias finitas y elementos finitos clásicos lineales, y también los esquemas SIPG en los cuales los números de ondas patológicos no se han filtrado previamente, se tiene que $v = 0$. Bajo un mecanismo de filtrado, v es el mínimo en el conjunto restringido de números de onda de la velocidad de grupo correspondiente.

En la Figura 1.1, observamos que la velocidad de grupo para la semi-discretización por diferencias finitas es estrictamente decreciente en $(0, \pi/h)$. En cambio, la velocidad de grupo correspondiente a la relación de dispersión física para la aproximación GD normalmente tiene un punto de máximo $\xi_s^* \in (0, \pi/h)$, de modo que crece en $(0, \xi_s^*)$ y decrece en $(\xi_s^*, \pi/h)$. Denotando por $\xi_s \in (\pi/2h, \pi/h)$ el número de onda en el que la velocidad de grupo física alcanza el valor uno, se observa que ξ_∞ corresponde al número de onda para el cual la velocidad de grupo correspondiente al método de elementos finitos clásicos lineales alcanza el valor uno. Para $s > 3$, se tiene que $\xi_s > \xi_\infty > \pi/2h$. Denotamos por $T = 2$ el tiempo de observabilidad minimal para la ecuación de ondas continua y por T_d^δ, T_∞^δ y T_s^δ los tiempos obtenidos para los esquemas en diferencias finitas, elementos finitos clásicos lineales y el de la rama física en SIPG, después de haber truncado las altas frecuencias en Fourier con un parámetro $\delta \in (0, 1)$. El algoritmo bimalla corresponde a $\delta = 1/2$. Entonces $T_s^{1/2} = T_\infty^{1/2} = T$, mientras que $T_d^{1/2} = 2\sqrt{2}$. Esto quiere decir que la proyección sobre la rama física en la semi-discretización por SIPG o la semi-discretización por elementos finitos clásicos lineales combinada con un algoritmo bimalla proporciona resultados óptimos desde el punto de vista de la propagación, mientras que el esquema en diferencias finitas combinado con el algoritmo bimalla produce un tiempo de observabilidad más largo. Cuando $\delta \in (\xi_s h/\pi, 1)$, ocurren las siguientes desigualdades:

$$T < T_s^\delta < T_\infty^\delta < T_d^\delta = \frac{2}{\cos(\pi\delta/2)}.$$

Esto muestra que los métodos SIPG, gracias a su mayor complejidad, proporcionan mejores resultados que las diferencias finitas o los elementos finitos clásicos lineales en las frecuencias más altas.

Esto significa que para todo $\delta \in (0, 1)$, $T > T_s^\delta$ y $\overrightarrow{U}^i = (\overrightarrow{A}^i, \overrightarrow{J}^i)$, con $\overrightarrow{J}^i = 0$ y $\overrightarrow{A}^i$ obtenidos al truncar en el espacio de Fourier por un parámetro δ, $i = 0, 1$, existe una sucesión de controles uniformemente acotada en h, $\overrightarrow{F}(t) \in L^2(0, T, (\ell^2(\Omega))^2)$ de modo que la solución del siguiente problema no-homogéneo

$$\begin{cases} M_h \overrightarrow{U}_{tt}(t) + R_h^s \overrightarrow{U}(t) = \overrightarrow{F}(t), & t \in (0, T) \\ \overrightarrow{U}(0) = \overrightarrow{U}^0, \quad \overrightarrow{U}_t(0) = \overrightarrow{U}^1 \end{cases}$$

verifica $\Gamma_h^\delta \overrightarrow{U}(T) = \Gamma_h^\delta \overrightarrow{U}_t(T) = 0$, donde Γ_h^δ es el operador de proyección definido como

$$\Gamma_h^\delta f_j = \frac{1}{2\pi} \int_{\Pi_h^\delta} \widehat{f}^h(\xi) \exp(i\xi x_j)\, d\xi,$$

para todo $\overrightarrow{f} \in \ell^2(\mathbb{R})$, y siendo $\widehat{f^h}$ la TFSD de $\overrightarrow{f}$.

Un valor interesante del parámetro de estabilización en los métodos SIPG es $s = 3$. En este caso, no existen los puntos críticos en $\xi = \pi/h$ en ninguna de las dos ramas de dispersión. Por tanto, sólo tomando datos iniciales con saltos nulos, se tiene que el tiempo de observabilidad correspondiente a la semi-discretización SIPG es el del caso continuo, $T_3^1 = 2$. Resulta que, a pesar de la complejidad del esquema numérico, para este valor particular del parámetro de estabilización s, el mecanismo de filtrado es más sencillo que el mecanismo de filtrado típico en los esquemas SIPG. Este valor aislado del parámetro de estabilización $s = 3$ sigue un fenómeno conocido sobre la discretización total de la ecuación de ondas $1-d$ por el esquema centrado en diferencias finitas. Se sabe que cuando los pasos espacial y temporal son iguales, la solución discreta coincide con la solución continua el los nodos del mallado (cf. [44]) y por tanto no hay altas frecuencias patológicas. Probablemente no tiene un análogo para los métodos SIPG de orden superior, en varias dimensiones espaciales sobre mallas uniformes y para aproximaciones SIPG completamente discretas de la ecuación de ondas.

En la última sección del capítulo, hacemos un análisis de Fourier riguroso de los otros métodos GD simétricos analizados de manera unitaria en [6]. En este marco simplificado $1-d$, concluimos que hay solo dos clases de métodos GD: el SIPG y el *Local Discontinuous Galerkin* (LDG) introducido y analizado en [16],[22]. Los métodos LDG son más complejos, en el sentido de que dependen de dos parámetros de estabilización: $s > 0$ y $\beta \in \mathbb{R}$. Los diagramas de dispersión revelan fenómenos más complejos en comparación con los métodos SIPG. Por ejemplo, para β grandes y s pequeños, la rama física tiene dos puntos críticos y puede generar incluso paquetes de onda que se propagan en la dirección no-física.

CAPÍTULO 3. Semi-discretizaciones por funciones spline de la ecuación de ondas $1-d$. En este capítulo analizamos y comparamos las propiedades de propagación para una clase de semi-discretizaciones de la ecuación de ondas $1-d$ que dependen de un parámetro $k \in \mathbb{N} \setminus \{0,1\}$, que denota la regularidad de las soluciones aproximadas. Es decir, las soluciones numéricas pertenecen al espacio $C^k(\mathbb{R})$. Se usa sólo un tipo de funciones de base soportadas en intervalos de longitud $(k+2)h$. Estos métodos se llaman en la literatura k-refinamientos (cf. [29]) o cardinal splines ([49], [20]). A pesar de su mayor regularidad y orden de aproximación de la ecuación de ondas continua en comparación con los elementos finitos clásicos lineales, estos métodos mantienen un comportamiento patológico a altas frecuencias. Más precisamente, producen sólo una relación de dispersión que tiene un único punto singular localizado en el número de onda $\xi = \pi/h$. Por tanto, se pueden construir paquetes de ondas patológicos en las más altas frecuencias siguiendo los métodos de los capítulos anteriores y, sin considerar datos iniciales filtrados, la constante de observabilidad asociada explota al menos a velocidad polinomial de orden arbitrario.

El caso $k = 0$ coincide con el esquema de elementos clásicos lineales, y en el límite cuando $k \to \infty$ se recupera la ecuación de ondas continua. Por consiguiente, a medida que k aumenta, la relación de dispersión correspondiente a un método C^k se acerca a la continua. Como se puede observar en la Figura 1.2, la velocidad de grupo correspondiente tiene un punto de máximo localizado en $(0, \pi/h)$ y alcanza por la segunda vez el valor uno en un número de onda $\xi_k \in (\pi/2h, \pi/h)$, donde ξ_k es creciente como función de k. Denotamos por T_k^δ el tiempo de observabilidad óptimo derivado del análisis del diagrama de dispersión cuando el dato inicial se trunca en Fourier con parámetro $\delta \in (0,1)$. Cuando los datos iniciales en el k-método están filtrados por un algoritmo bimalla, se observa que el tiempo de observabilidad es óptimo, $T_k^{1/2} = 2$, para todo $k \geq 0$. En cambio, cuando los datos iniciales se filtran por un procedimiento de truncamiento en el espacio de Fourier con un parámetro $\delta \in (\xi_{k'} h/\pi, 1)$, entonces $2 < T_{k'}^\delta < T_k^\delta$, para todo $0 \leq k < k'$, lo que muestra que los resultados de estos métodos mejoran en las frecuancias más altas a medida que la regularidad k aumenta.

CAPÍTULO 4. Semi-discretizaciones por elementos finitos clásicos de orden superior de la ecuación de ondas $1-d$. En este capítulo consideramos una clase de aproximaciones numéricas más sofisticadas de la ecuación de ondas $1-d$ generados por semi-discretizaciones en

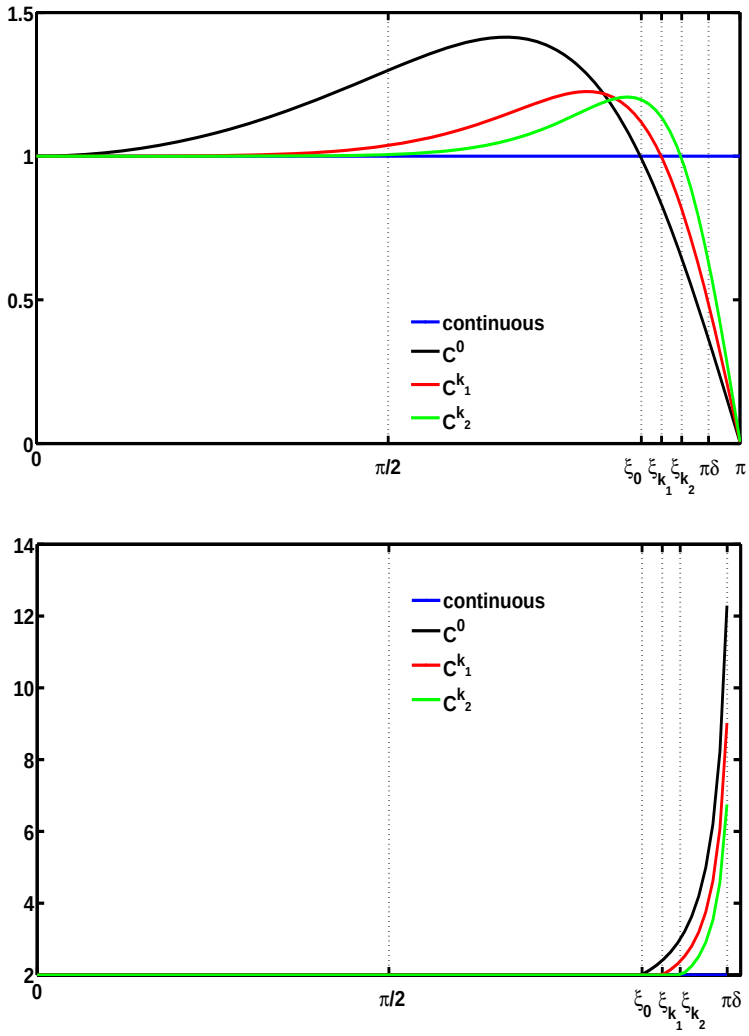

Figure 1.2: Velocidad de grupo versus tiempo de observabilidad óptimo correspondientes a los métodos C^k, $k = 0$, $k_1 = 1$ y $k_2 = 2$.

espacio por elementos finitos clásicos de orden superior en una malla uniforme de paso $h > 0$. En cada celda computacional (x_j, x_{j+1}), se añade una malla auxiliar, llamada *serendipity mesh* (cf. [11]), formada por los puntos $x_j + l\widetilde{h}$, $l = 1, \ldots, k-1$, $\widetilde{h} = h/k$. Con respeto al capítulo anterior, la aproximación por elementos finitos clásicos P_k que usamos en este capítulo usa k tipos de funciones de base que pertenecen al espacio $C^0(\mathbb{R})$. Una representa los puntos nodales x_j, es soportada en (x_{j-1}, x_{j+1}) y toma el valor uno en x_j y cero en todos los *serendipity points* de $[x_{j-1}, x_{j+1}]$. Las restantes $k-1$ funciones están soportadas en $[x_j, x_{j+1}]$ y toman el valor uno en uno de los *serendipity points* y cero en los demás.

Para simplificar, consideramos sólo aproximaciones cuadráticas. El motivo para hacer esta restricción es que usamos de manera explícita las expresiones de los autovaleres del símbolo de Fourier, que para el caso general de los polinomios de orden k es una matriz $k \times k$. Es técnicamente complicado resolver exactamente una ecuación característica de orden k para encontrar los autovalores que dependen además del número de onda ξ. Por esto, nos restringimos al caso particulare $k = 2$, para el cual existen fórmulas explícitas para encontrar las raíces de un polinomio de orden

k. No obstante, un análisis riguroso de estos casos particulares indica cómo debe procederse en el caso de un k general.

A pesar de que, cuando los datos iniciales en el modelo continuo son de clase $C^\infty(\mathbb{R})$ el rendimiento de estas aproximaciones aumenta con el k, aparecen, sin ambargo, ciertos fenómenos singulares en las propiedades de propagación. En concreto, la semi-discretización por elementos P_k de la ecuación de ondas produce k diagramas de dispersión (una acústica y $k-1$ ópticas). En el diagrama acústico, existe un único punto crítico ubicado en el número de onda $\xi = \pi/h$, y en cada uno de los diagramas de dispersión ópticos hay dos puntos singulares ubicados en los números de onda $\xi = 0$ y $\xi = \pi/h$. Hemos analizado este tipo de comportamiento en el segundo capítulo, para la semidiscretización por métodos GD P_1. Por tanto, adaptando las construcciones previas de paquetes de ondas que se propagan a velocidad arbitrariamente lenta, se puede concluir que la constante de observabilidad correspondiente a la semi-discretización por elementos finitos cuadráticos explota también de manera polinomial de orden arbitrario.

El primer mecanismo de filtrado que desarrollamos para elementos finitos clásicos de orden k sigue las líneas del primero que habíamos descrito para los métodos GD, y consiste en considerar soluciones que involucren sólo la rama acústica de frecuencias. Para eliminar la patología generada por el punto crítico existente en dicha rama en $\xi = \pi/h$, se usa o bien un truncamiento en el espacio de Fourier, o bien un algoritmo bimalla.

El mecanismo de filtrado B) descrito para las aproximaciones GD se basaba en la idea de que eliminar los saltos de los datos iniciales numéricos debería ayudar a eliminar la patología generada por la singularidad en la rama espúrea en el número de onda $\xi = 0$. El análogo de esto para la aproximación por elementos finitos clásicos de orden superior considerada a lo largo de este capítulo consiste en considerar datos iniciales numéricos lineales. Para la aproximación por elementos finitos cuadráticos, esto significa que los valores de los datos iniciales en los puntos medios son una media de los correspondientes valores en los dos nodos vecinos. Una vez elegidos los datos iniciales de esta manera, de modo que la mayor parte de su energía se concentre en la rama acústica, las componentes espúreas en la alta frecuencia π/h se pueden filtrar mediante una truncamiento en el espacio de Fourier o un algoritmo bimalla.

Para $k = 2$, la relación de dispersión acústica resultante aproxima mejor la continua que la correspondiente a elementos finitos clásicos lineales. La velocidad de grupo correspondiente a la rama acústica tiene un punto de máximo ubicado en $\xi_k^*(0, \pi/h)$. Denotamos por $\xi_k \in (\xi_k^*, \pi/h)$ el número de onda donde esta velocidad de grupo alcanza el valor uno. Se tiene que $\xi_k > \pi/2h$. Denotamos por T_k^δ el tiempo de observabilidad uniforme obtenido al analizar la rama de dispersión acústica cuando los datos iniciales se truncan en el espacio de Fourier por el parámetro $\delta \in (0, 1)$. Considerando datos iniciales lineales con valores nodales dados por un algoritmo bimalla, se obtiene que el tiempo de observabilidad es el óptimo, $T_k^{1/2} = 2$. Sin embargo, usando un truncamiento en Fourier con $\delta \in (\xi_{k'} h/\pi, 1)$, el tiempo de observabilidad $T_{k'}^\delta$ verifica $2 < T_{k'}^\delta < T_k^\delta$, para todo $k < k'$. Esto hace que el rendimiento de los elementos finitos clásicos de orden k aumente con k en las más altas frecuencias.

CAPÍTULO 5. Conclusiones y problemas abiertos. Acabamos esta memoria con un listado de problemas abiertos y líneas de investigación futuras relacionados con los temas que hemos desarrollado a lo largo de la presente tesis.

Chapter 2

Introduction

This thesis is devoted to analyze the consequences of increasing the complexity of numerical schemes on the approximation of wave propagation and dispersion phenomena. As it is well known, convergent numerical methods not only reproduce the true solutions of the continuous models as Partial Differential Equations (PDEs), but may also generate spurious high frequency numerical artifacts. This fact is not an impediment for the convergence property to hold in the classical sense of numerical analysis, in the presence of well known fixed initial and boundary data, but may be a major drawback when one is dealing with large classes of data and solutions, as it happens in the frame of optimal design, inverse and control problems. In these contexts, it is well known that the classical requirements of consistency and stability that guarantee the convergence of approximations of PDEs do not suffice. For instance, in [59] it was proved that the control for a finite difference approximation of the $1-d$ wave equation may diverge as the mesh size tends to zero, even if the limiting wave equation is controllable. Similarly, in [31], it was proved that the classical dispersive Strichartz inequalities fail to be true uniformly with respect to the mesh size parameter for the numerical approximations of the Schrödinger equation. The first main contribution of this thesis is to make a systematic analysis of the high frequency pathologies that are responsible for those instabilities.

We will rigorously build high frequency solutions localized both in the physical and Fourier space so that their support remains localized along characteristic rays propagating with the so-called *group velocity* of the numerical scheme, which may degenerate as the mesh size tends to zero for typical numerical approximation schemes of the wave equation. This is in contrast with the fact that all solutions propagate with an uniform velocity for the corresponding continuous models. We develop several high-frequency constructions yielding to similar results based on: plane wave analysis, Gaussian beams or WKB expansions. These constructions lead to spurious high-frequency wave packets. Our analysis makes more precise the Fourier analysis in N. Trefethen (cf. [54]) and later works by L. Ignat, M. Negreanu and E. Zuazua (see [32], [31], [45], among others). We first develop it in the context of finite-differences, to later consider more sophisticated numerical methods based on the discontinuous Galerkin (DG) approximations or higher order finite elements. As we shall see, even increasing the approximation order of the numerical schemes or by relaxing them through a DG approach, these high frequency pathologies remain. The DG approximations are more complex, in the sense that they involve several dispersion relations. Each of these dispersion relations, a physical one and at least a spurious one, has its own critical points. In view of this, using decoupling arguments in the Fourier space, one can construct high-frequency wave packets around each of these critical points of the spectrum, denying any type of uniform propagation or dispersion property. We show some interesting fine properties of these wave packets in what concerns their dependence with respect to the mesh size parameter as, for example, the divergence rate of the associated observability constant or their lack of symmetry. For the DG methods or higher order classical finite element approximations, there exist even wave packets propagating in the wrong direction located on the spurious diagrams.

In view of these pathologies, in the literature devoted to develop stable numerical algorithms for the control of waves and solving dispersive equations, several devices has been developed to attenuate the effect of the high frequency components by considering, for instance, suitable classes of filtered initial data using Fourier truncation of the high frequencies (cf. [58]), mixed finite elements methods (cf. [17]), two-grid algorithms ([32], [45]), vanishing numerical viscosity (cf. [52]) and Tychonoff regularization (cf. [26]). In this thesis, we also adapt these filtering mechanisms to the case of the DG or higher order classical finite element approximations. Due to the more complex behavior of these schemes, a careful combination of several filtering techniques is required to effectively eliminate the spurious effects.

To be more precise, in this work, we address the following four topics:

1. **Concentrated waves for the finite difference approximations of the wave equation.** We develop rigorous constructions of high frequency wave packets for the finite difference semi-discretization of the wave equation by several techniques (asymptotic methods - WKB or Gaussian beams - and the construction of explicit exponentially concentrated data in the semi-discrete wave equation). In particular, we prove the lack of uniform observability properties for the semi-discrete wave equation and the polynomial blow-up (with an arbitrarily large rate) of the observability constant.

2. **Propagation properties for discontinuous Galerkin approximations of the $1-d$ wave equation.** We make a systematic study of the propagation properties of the symmetric discontinuous Galerkin (DG) discretizations of the $1-d$ wave equations. More precisely, we make a careful Fourier analysis of some classes of DG approximations on uniform grids and construct high-frequency concentrated numerical solutions, showing the lack of uniform observability properties for the wave equation. We also design filtering mechanisms for the semi-discrete wave equation based on Fourier truncation and bi-grid algorithm.

3. **Propagation properties for the spline approximations of the wave equation.** We develop a systematic analysis of the propagation properties of the so-called k-refinements or C^k methods consisting in using a basis of spline functions to approximate the solutions of the $1-d$ wave equation. We show that for any degree of regularity k of these methods, the velocity of propagation of the wave packets still has not a strictly positive lower bound. Nevertheless, as k increases, the corresponding dispersion relation fits better to the continuous one for a larger range of wave numbers than the one corresponding to the linear finite element approximation.

4. **Propagation properties for the quadratic classical finite element semi-discretization of the wave equation.** We perform a rigorous Fourier analysis of the quadratic classical finite element approximation of the wave equation. Similar phenomena to the DG semi-discretization appear. We also develop the corresponding filtering mechanisms.

At the end of the thesis, we present a list of open problems related to the topics we address within the Thesis that may constitute interesting future research lines.

In the following, we briefly describe the problems under consideration and the main results of the thesis, underlying also the main novelties of the techniques we develop.

CHAPTER 1. Finite difference approximations of the wave equation. The problem of constructing wave packets propagating with the group velocity for numerical discretizations of wave or transport equations has been addressed before, starting with the pioneering paper due to Trefethen (cf. [54]). Nevertheless, our analysis is deeper and yields more precise results. Our constructions are motivated by control theoretical issues, but they are also of independent interest.

Given $T > 0$ and a mesh size $h > 0$, we consider the Cauchy problem associated to the semi-discrete wave equation

$$\begin{cases} \partial_t^2 u_{\mathbf{j}}(t) - \Delta_h u_{\mathbf{j}}(t) = 0, & \mathbf{j} \in \mathbb{Z}^d,\ t \in (0,T) \\ u_{\mathbf{j}}(0) = u_{\mathbf{j}}^0,\ \partial_t u_{\mathbf{j}}(0) = u_{\mathbf{j}}^1, & \mathbf{j} \in \mathbb{Z}^d, \end{cases} \tag{2.1}$$

with $\Delta_h u_{\mathbf{j}} = h^{-2} \sum_{k=1}^{d} \big(u_{\mathbf{j}+\mathbf{e}_k} - 2u_{\mathbf{j}} + u_{\mathbf{j}-\mathbf{e}_k}\big)$.

This constitutes a semi-discrete finite difference approximation of the Cauchy problem for the wave equation

$$\begin{cases} \partial_t^2 u(x,t) - \Delta u(x,t) = 0, & x \in \mathbb{R}^d,\ t \in (0,T) \\ u(x,0) = u^0(x), \quad \partial_t u(x,0) = u^1(x), & x \in \mathbb{R}^d \end{cases} \tag{2.2}$$

on an uniform space grid of mesh-size h. Obviously, $u_{\mathbf{j}}(t)$ stands for an approximation of $u(x,t)$ at the point $x = x_{\mathbf{j}} = \mathbf{j}h$, $\mathbf{j} \in \mathbb{Z}^d$.

We also introduce the discrete gradient $\nabla_h = (\partial_{h,k})_{k=1,\dots,d}$, where

$$\partial_{h,k} \overrightarrow{f} = \frac{\overrightarrow{f}_{\cdot+\mathbf{e}_k} - \overrightarrow{f}}{h},$$

$\overrightarrow{f} = (f_{\mathbf{j}})_{\mathbf{j}\in\mathbb{Z}^d}$ is any sequence and $(\mathbf{e}_l)_{l=1}^d$ is the canonical basis in $\mathbb{R}^d$. For an open set $O \subset \mathbb{R}^d$, we define the discrete Sobolev spaces

$$\ell^2(O) = \{\overrightarrow{f} \text{ s.t. } \|\overrightarrow{f}\|^2_{\ell^2(O)} := h^d \sum_{x_{\mathbf{j}}\in O} |f_{\mathbf{j}}|^2 < \infty\}$$

and

$$\dot{h}^1(O) = \{\overrightarrow{f} \text{ s.t. } \|\overrightarrow{f}\|^2_{\dot{h}^1(O)} := \sum_{k=1}^{d} \|\partial_{h,k}\overrightarrow{f}\|^2_{\ell^2(O)} < \infty\}.$$

If $\Gamma \subset O$ is an hyper-surface, then the following discrete Sobolev space on Γ is defined:

$$\ell^2(\Gamma) = \{\overrightarrow{f} = (f_{\mathbf{j}})_{x_{\mathbf{j}}\in\Gamma} : \|\overrightarrow{f}\|^2_{\ell^2(\gamma)} := h^{d-1} \sum_{x_{\mathbf{j}}\in\Gamma} |f_{\mathbf{j}}|^2 < \infty\}.$$

Set $\Omega := \mathbb{R}^d \setminus B^d(0,1)$, where $B^d(0,1)$ is the d-dimensional unit ball. Our original motivation is the behavior as $h \to 0$ of the constant:

$$C_h(T) = \sup_{(\overrightarrow{u}^0,\overrightarrow{u}^1)\in\dot{h}^1(\mathbb{R}^d)\times\ell^2(\mathbb{R}^d)} \frac{\|\overrightarrow{u}^0\|^2_{\dot{h}^1(\mathbb{R}^d)} + \|\overrightarrow{u}^1\|^2_{\ell^2(\mathbb{R}^d)}}{\int_0^T (\|\nabla_h \overrightarrow{u}(t)\|^2_{\ell^2(\Omega)} + \|\partial_t \overrightarrow{u}(t)\|^2_{\ell^2(\Omega)})\, dt}. \tag{2.3}$$

Whatever $T > 0$ and $h > 0$ are, this constant measures how efficiently the total energy of the solutions of (2.1) can be observed by making measurements in the outer region $|x| > 1$ during the finite time interval $t \in (0,T)$ and it can be shown to be finite. However, it is well-known to diverge as $h \to 0$ whatever $T > 0$ is (cf. [59]). Our goal here is to analyze the rate of divergence. Note that the constant $C_h(T)$, the so-called *observability constant*, is well-known to be finite for the continuous wave equation (2.2) for $T > 2$ ([7], [36]). However, for the numerical approximations, the observability constant $C_h(T)$ blows-up as $h \to 0$, whatever $T > 0$ is, which shows the instability of the numerical scheme in a control theoretical context.

The observability problem we study in this chapter is interesting by itself, but it has some applications in control theory, inverse and stabilization problems. In particular, problem (2.1) can be shown to be equivalent (by means of the Hilbert Uniqueness Method (HUM) introduced in [36]) to an exact controllability problem through a control supported in the set $\Omega = \mathbb{R}^d \setminus B^d(0,1)$. More precisely, the corresponding control problem can be formulated as follows. For all $(\overrightarrow{u}^0, \overrightarrow{u}^1) \in \dot{h}^1(h\mathbb{Z}^d) \times \ell^2(h\mathbb{Z}^d)$ and all finite time $T > 0$, find a control $(f_{\mathbf{j}}\chi_{|x_{\mathbf{j}}|>1}(\mathbf{j}))_{\mathbf{j}\in\mathbb{Z}^d}$ such that the solution to the problem

$$\begin{cases} \partial_t^2 u_{\mathbf{j}}(t) - \Delta_h u_{\mathbf{j}}(t) = f_{\mathbf{j}}\chi_{|x_{\mathbf{j}}|>1}(\mathbf{j}), & \mathbf{j} \in \mathbb{Z}^d,\ t \in (0,T) \\ u_{\mathbf{j}}(0) = u^0_{\mathbf{j}},\ \partial_t u_{\mathbf{j}}(0) = u^1_{\mathbf{j}}, & \mathbf{j} \in \mathbb{Z}^d, \end{cases} \tag{2.4}$$

satisfies $u_{\mathbf{j}}(T) = \partial_t u_{\mathbf{j}}(T) = 0$. The fact that the observability constant $C_h(T)$ diverges polynomially implies that there exist bounded classes of initial data in (2.4) for which the minimal control f in $L^2(0,T,\ell^2(\Omega))$ diverges at least polynomially at arbitrary order in $L^2(0,T;\ell^2(\Omega))$ as $h \to 0$.

Similar control problems can be analyzed in bounded domains with interior or boundary control. More precisely, for $h = 1/(N+1)$ and $\Gamma_0 \subset \Gamma = \partial\mathcal{I}$, $\mathcal{I} := (0,1)^d$, consider the semi-discrete boundary initial value problem with boundary control

$$\begin{cases} \partial_t^2 u_{\mathbf{j}}(t) - \Delta_h u_{\mathbf{j}}(t) = 0, & x_{\mathbf{j}} \in \mathcal{I},\ t \in (0,T) \\ u_{\mathbf{j}}(t) = v_{\mathbf{j}}^h(t)\chi_{\Gamma_0}(\mathbf{j}), & x_{\mathbf{j}} \in \Gamma,\ t \in (0,T) \\ u_{\mathbf{j}}(0) = u_{\mathbf{j}}^0,\ \partial_t u_{\mathbf{j}}(0) = u_{\mathbf{j}}^1, & x_{\mathbf{j}} \in \mathcal{I}. \end{cases} \tag{2.5}$$

The exact controllability problem requires to find the control $\overrightarrow{v}^h \in L^2(0,T;\ell^2(\Gamma_0))$ such that for all $(\overrightarrow{u}^0, \overrightarrow{u}^1) \in \ell^2(\mathcal{I}) \times \hbar^{-1}(\mathcal{I})$ the solution of (2.5) satisfies $u_{\mathbf{j}}(T) = \partial_t u_{\mathbf{j}}(T) = 0$, for all $x_{\mathbf{j}} \in \mathcal{I}$. Here $\hbar^{-1}(\mathcal{I})$ is the dual of $\hbar_0^1(\mathcal{I})$, the discrete analogous of $H_0^1(\mathcal{I})$, defined by

$$\hbar_0^1(\mathcal{I}) = \{\overrightarrow{f} = (f_{\mathbf{j}})_{x_{\mathbf{j}} \in \mathcal{I}} : \|\overrightarrow{f}\|_{\hbar_0^1(\mathcal{I})} := \|\overrightarrow{f}\|_{\dot{\hbar}^1(\mathcal{I})} < \infty, f_{\mathbf{j}} = 0, x_{\mathbf{j}} \in \Gamma\}.$$

By the Hilbert Uniqueness Method (HUM), the exact controllability problem (2.5) is equivalent to an observability problem for the semi-discrete adjoint problem,

$$\begin{cases} \partial_t^2 \phi_{\mathbf{j}}(t) - \Delta_h \phi_{\mathbf{j}}(t) = 0, & x_{\mathbf{j}} \in \mathcal{I},\ t \in (0,T) \\ \phi_{\mathbf{j}}(t) = 0, & x_{\mathbf{j}} \in \Gamma,\ t \in (0,T) \\ \phi_{\mathbf{j}}(0) = \phi_{\mathbf{j}}^0,\ \partial_t \phi_{\mathbf{j}}(0) = \phi_{\mathbf{j}}^1, & x_{\mathbf{j}} \in \mathcal{I}. \end{cases} \tag{2.6}$$

More precisely, the exact controllability property holds if and only if the following observability constant is finite:

$$C_h(T) = \sup_{(\overrightarrow{\phi}^0, \overrightarrow{\phi}^1) \in \hbar_0^1(\mathcal{I}) \times \ell^2(\mathcal{I})} \frac{\|\overrightarrow{\phi}^0\|_{\hbar_0^1(\mathcal{I})}^2 + \|\overrightarrow{\phi}^1\|_{\ell^2(\mathcal{I})}^2}{\sum\limits_{\mathbf{j} \in \Gamma_{h,0}} \int\limits_0^T \|\partial_{h,\nu} \overrightarrow{\phi}(t)\|_{\ell^2(\Gamma_0)}^2 \, dt}, \tag{2.7}$$

where the supremum is taken over all the solutions $\overrightarrow{\phi}$ of finite energy of (2.6), ν is the outward unit normal vector to Γ_0 and $\partial_{h,\nu}$ is the discrete normal derivative defined by

$$\partial_{h,\nu} f_{\mathbf{j}} = -\frac{1}{h} f_{\mathbf{j}-\nu}, \forall x_{\mathbf{j}} \in \Gamma_0,\ \overrightarrow{f} \in \hbar_0^1(\mathcal{I}).$$

In [33] or [58], it has been shown that, when $d = 1$ or $d = 2$, $C_h(T) \to \infty$ as $h \to 0$ since the *gap* of the eigenvalues located in small neighborhoods of the wave numbers $\{0, \pi/h\}^d \setminus \{0\}$ tends to zero as $h \to 0$. Also the Fourier truncation has been shown to be effective to recover the uniform observability property in suitable classes of low-frequency solutions. In [45] and [32] a bi-grid algorithm for the $1-d$ and $2-d$ versions of (2.6) respectively has been developed and its efficiency to reestablish the uniformity of the constant $C_h(T)$ as $h \to 0$ has been shown. In [42] a fine analysis of the controls in (2.5) is carried out for the $1-d$ case. Using a fine analysis of bi-orthogonal sequences, Micu shows the existence of a control of size $\exp(\sqrt{N})$. By HUM, this means that $C_h(T)$ defined by (2.7) increases at least exponentially. Using this result in [42], we show in Proposition 1.2.1, pp. 7, the existence of wave packets for which the observability constant $C_h(T)$ in (2.3) blows-up exponentially in $d = 1$.

One of our original contributions is to find a lower bound for the divergence rate of $C_h(T)$ which is valid in any space dimension, for the Cauchy problem in the whole space. More precisely, we prove that $C_h(T)$ diverges polynomially at any order by developing several slightly different constructions of high frequency wave packets. Our result generalizes that in [42] showing the exponential divergence of the observability constant in $1-d$ and with Dirichlet boundary conditions, although we do not get an exponential divergence rate. The high frequency constructions we develop are as follows:

- Explicit exact solutions given smooth initial data belonging to the Schwartz class with support of order h^α in the physical space, for some $\alpha < 1$.
- Discrete WKB expansions.
- Discrete *Gaussian beams* constructions, following the previous construction in the continuous case due to Ralston, [46].

Although the solutions we build solve the Cauchy problem (2.1), in view of their localization in space-time, they can be easily used to prove also the polynomial divergence (at any order) of the observability constants for discrete wave problems in bounded domains with various boundary conditions.

We also present several plots of these high-frequency wave packets and make a careful analysis of some of their qualitative properties, especially the dispersive ones. Thus, we observe and show that, although the initial data are symmetric, for example samples of Gaussian functions, the numerical solution at future times does not maintain this symmetry. This is due to the fact that the wave packets we construct travel along characteristics, but not all the wave fronts travel at the same velocity. This breaks the symmetry existing in the continuous solution corresponding to the same initial data. Another property that we analyze is their likeness to a Gaussian function. We show that for $\alpha < 2/3$, for finite times, the wave packets we construct can be approximated by a Gaussian profile. The numerical simulations show that for larger values of α this similarity to a Gaussian profile is loosed, but, so far, we do not have a rigorous proof of this fact.

The construction in this chapter will be adapted within next chapters to more complex numerical schemes, in particular those arising from DG and higher order classical FEM approximations of the $1-d$ wave equation.

CHAPTER 2. Discontinuous Galerkin semi-discretizations of the $1-d$ wave equation. The discontinuous Galerkin (DG) methods are non-conforming approximations of PDEs, yielding solutions that present discontinuities along the interfaces. In order to ensure the stability of this kind of approximations, one has to penalize the jumps by adding numerical fluxes depending on some stability parameters.

In the second chapter of this Thesis, we analyze the propagation properties for some DG semi-discretizations of the $1-d$ wave equation. The simplest case of P_1 polynomials on an uniform grid of size h is considered. Typically, a semi-discretization of this type can be written in the form

$$M_h \overrightarrow{U}_{tt}(t) + R_h^s \overrightarrow{U}(t) = 0,$$

where $\overrightarrow{U}(t) = (A_k(t), J_k(t))_{k\in\mathbb{Z}}$ is the vector of unknowns constituted by two kinds of components, $A_k(t)$ and $J_k(t)$, representing the average and the jump of the numerical solution at the grid point x_k, $k \in \mathbb{Z}$. Here and in the sequel, R_h^s and M_h are the stiffness and the mass block-three-diagonal infinite matrices, depending on h and on the stability parameter $s > 1$.

Taking semi-discrete Fourier transforms (SDFTs) in the space variable, the above infinite system of ODEs can be transformed into a 2×2-system of ODEs depending on the parameter ξ:

$$\widehat{U}_{tt}^h(\xi, t) + S_h^s(\xi)\widehat{U}^h(\xi, t) = 0, \tag{2.8}$$

where $S_h^s(\xi) = (M_h(\xi))^{-1}R_h^s(\xi)$ and $M_h(\xi), R_h^s(\xi)$ are the 2×2 matrix symbols of M_h and R_h^s depending on the Fourier variable $\xi \in \Pi_h$. The two components of the unknown vector $\widehat{U}^h(\xi, t)$ represent the continuous and the jump part of the numerical solution, respectively. Clearly, in order to obtain the uniqueness of the solution, one has to complement the system (2.8) with two initial data, $\widehat{U}^h(\xi, 0)$ and $\widehat{U}_t^h(\xi, 0)$.

For the *Symmetric Interior Penalty Discontinuous Galerkin* method (SIPG) (cf. [5]), the Fourier analysis we develop demonstrates the coexistence of two dispersion diagrams, which are in fact the square roots of the two eigenvalues of the matrix $S_h^s(\xi)$: a physical (acoustical) one,

which is the smallest of the two branches and whose behavior is similar to the dispersion relation for the P_1 classical FEM, and a spurious (optical) one, which tends to infinity as h goes to zero or the stabilization parameter s tends to infinity.

Although the physical dispersion diagram for the SIPG methods is much closer to the continuous one than when using standard finite differences or P_1-finite element methods, the group velocity analysis shows that both acoustical and optical branches present singular points: one located on the physical dispersion relation at the wave number π/h and at least two other ones located on the spurious diagram at the wave number $\xi = 0$ and $\xi = \pi/h$. Having singular points on the dispersion diagrams at the highest wave number π/h is a classical fact, well-known for the finite difference and finite element schemes. The singularity at the wave number 0 on the optical diagram is less usual. This is due to the symmetry properties of the spurious diagram with respect to $\xi = 0$, since the dispersion relation is a rational expression of trigonometric polynomials, and to the fact that the corresponding Fourier symbol of the mass matrix $M_h(\xi)$ is not singular.

For each one of these singular points, the constructions of concentrated solutions in Chapter 1 can be adapted. In order to build wave packets propagating with the group velocity corresponding to only one of the two dispersion diagrams and to avoid the interaction between the two modes, a decoupling algorithm in the Fourier space is used. This yields numerical solutions involving only one dispersion relation. Furthermore, one has to take into account the fact that, on the spurious diagram, in general, the wave packets travel in an unphysical direction, which make them doubly unphysical: in view of the propagation velocity and also of the direction of propagation.

We also design appropriate filtering mechanisms for this class of complex semi-discretizations to recover the uniformity of the observability constant with respect to the mesh size h in suitable subspaces of numerical solutions.

Two main strategies are analyzed:

A) By a decoupling argument in the Fourier space and an appropriate choice of the initial data, we deal with solutions of the SIPG semi-discretization involving only the physical dispersion relation having only one singular point. Then one can apply well known filtering algorithms for the initial data: *a) Fourier truncation* (initial data for which the support of the corresponding semi-discrete Fourier transforms does not contain the critical points of the corresponding dispersion relation, for instance $\xi \in [-\pi\delta/h, \pi\delta/h]$, with $\delta \in (0, 1)$) or *b) a bi-grid algorithm* (taking initial data obtained by linear interpolation from a coarser grid).

B) The second strategy is designed so that, given the initial data of the continuous model, the corresponding projections onto the numerical mesh are so that most of their energy is concentrated on the physical branch of the dispersion diagram. The simplest case of such a projection is when $J_k(0) = \partial_t J_k(0) = 0$, for all $k \in \mathbb{Z}$, i. e. when the numerical initial data are taken so that their jumps vanish. This will give rise to a weight for the spurious components in the expression of the SDFT of the solution vanishing at the wave number $\xi = 0$ and eliminating the pathological behavior at that wave number. In order to eliminate the pathological behavior at $\xi = \pi/h$, one has to apply an added *a)Fourier truncation* or *b)a bi-grid algorithm* on the physical initial data $(A_k(0), \partial_t A_k(0))_{k\in\mathbb{Z}}$.

From a practical viewpoint, the most feasible and useful filtering strategy consists in first choosing the jumps of the initial data to vanish and then to apply the bi-grid algorithm. This ensures that the energy of the numerical solutions is concentrated in the low frequencies of the physical branch. Furthermore, it has the advantage of being applicable on the physical grid without requiring the use of the SDFT.

The analysis developed in this chapter shows that the overall behavior of the numerical solutions obtained by the DG methods is more complex than that of most classical numerical methods such as finite differences or P_1 classical finite elements. But, after suitably filtering the initial data under consideration, the obtained solution turns out to be closer to those of the original wave equation than those that most classical methods yield. The observability time given by the *Geometric Control Condition* (cf. [7], which states that the observability time is the time needed for all

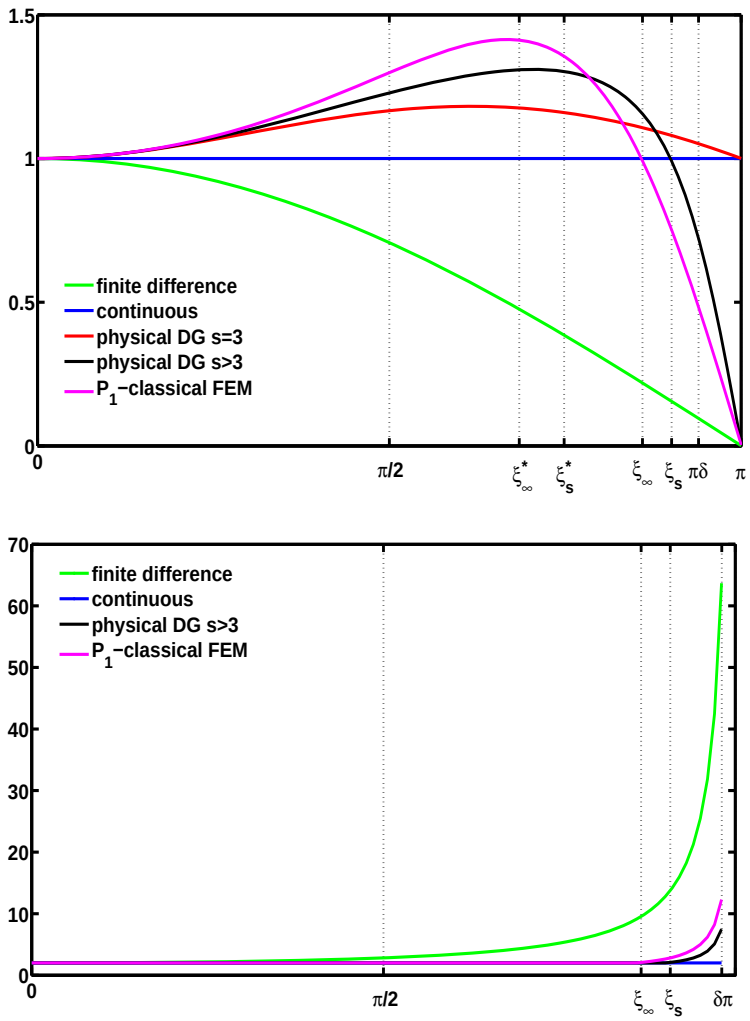

Figure 2.1: Group velocity versus optimal observability time corresponding to the finite difference, the physical mode in the SIPG method and linear classical finite element semi-discretizations of the $1-d$ wave equation.

the characteristics to enter the observation region) is $2/v$, where v is the minimal velocity of propagation of a Fourier component in the corresponding wave packet. For the most classical finite difference and classical finite element schemes and also for the SIPG schemes in which the pathological wave numbers have not been filtered out, one has $v = 0$. Under any kind of high frequency filtering, v is the minimum over the restricted set of wave numbers of the corresponding group velocity.

In Figure 2.1, we observe that the group velocity for the finite difference semi-discretization is strictly decreasing on $(0, \pi/h)$. Instead, the group velocity corresponding to the physical dispersion relation for the SIPG approximations typically has a maximum point $\xi_s^* \in (0, \pi/h)$, so that it increases in $(0, \xi_s^*)$ and decreases in $(\xi_s^*, \pi/h)$. Denote by $\xi_s \in (\pi/2h, \pi/h)$ the wave number where the group velocity reaches the value one and note that ξ_∞ corresponds to the wave number where the group velocity corresponding to the P_1-classical finite element method reaches the value one. For $s > 3$, we have that $\xi_s > \xi_\infty > \pi/2h$. Let us denote by $T = 2$ the minimal observability time

for the continuous wave equation and by T_d^δ, T_∞^δ and T_s^δ the ones obtained for the finite difference, linear classical finite element and the physical branch of the SIPG schemes after Fourier truncation with parameter $\delta \in (0,1)$. The bi-grid algorithm corresponds to $\delta = 1/2$. Then $T_s^{1/2} = T_\infty^{1/2} = T$, whereas $T_d^{1/2} = 2\sqrt{2}$, i.e. the projection of the solution on the physical branch in a SIPG method and the P_1-classical finite element semi-discretizations combined with a bi-grid algorithm give optimal results from a propagation viewpoint, whereas the finite difference scheme combined with a bi-grid algorithm gives a larger observability time. When $\delta \in (\xi_s h/\pi, 1)$, the following relation holds:

$$T < T_s^\delta < T_\infty^\delta < T_d^\delta = \frac{2}{\cos(\pi\delta/2)}.$$

This shows that the the DG method, despite its higher complexity, is the one yielding better results than the finite difference or the P_1-classical finite element method in the highest frequencies.

This means that for all $\delta \in (0,1)$, $T > T_s^\delta$ and $\overrightarrow{U}^i = (\overrightarrow{A}^i, \overrightarrow{J}^i)$, with $\overrightarrow{J}^i = 0$ and $\overrightarrow{A}^i$ obtained by a Fourier truncation of parameter δ, $i = 0, 1$, there exists an uniformly bounded (with respect to h) sequence of controls $\overrightarrow{F}(t) \in L^2(0, T, (\ell^2(\Omega))^2)$ such that the solution to the following inhomogeneous problem

$$\begin{cases} M_h \overrightarrow{U}_{tt}(t) + R_h^s \overrightarrow{U}(t) = \overrightarrow{F}(t), & t \in (0,T) \\ \overrightarrow{U}(0) = \overrightarrow{U}^0, \quad \overrightarrow{U}_t(0) = \overrightarrow{U}^1 \end{cases}$$

verifies $\Gamma_h^\delta \overrightarrow{U}(T) = \Gamma_h^\delta \overrightarrow{U}_t(T) = 0$, where Γ_h^δ is the projection operator defined as

$$\Gamma_h^\delta f_j = \frac{1}{2\pi} \int_{\Pi_h^\delta} \widehat{f^h}(\xi) \exp(i\xi x_j)\, d\xi,$$

for any $\overrightarrow{f} \in \ell^2(\mathbb{R})$, with $\widehat{f^h}$ being the SDFT of $\overrightarrow{f}$.

An interesting value of the stabilization parameter in the SIPG semi-discretization is $s = 3$. In that case, there are no critical points at $\xi = \pi/h$, neither on the physical branch, nor on the spurious one. Thus, only by taking initial data having no jumps, one has $T_3^1 = T$. It follows that, despite the complexity of the numerical scheme, for this particular value of the stabilization parameter s, the filtering mechanism is simpler than typically happens for the SIPG schemes. This isolated value of the stabilization parameter $s = 3$ follows the well-known fact on the fully-discrete finite difference scheme of the $1-d$ wave equation. Namely, when the time and space steps are equal, the discrete solution coincides with the continuous one in the grid points (cf. [44]) and there are not pathological high frequencies. Probably it has no analogous for higher order or higher space dimensions SIPG methods on uniform grids or for fully discrete SIPG schemes for the wave equation.

In the last section of this chapter, we make a rigorous Fourier analysis of the other symmetric DG methods analyzed in an unified framework in [6]. In this simplified $1-d$ framework, we conclude that there are only two important DG methods: the SIPG and the so-called Local Discontinuous Galerkin (LDG) method introduced and analyzed in [16],[22]. The LDG methods are more complex, in the sense that they depend on two stability parameters: $s > 0$ and $\beta \in \mathbb{R}$. The dispersion diagrams reveal more complex phenomena compared to the SIPG methods. For example, for large β and small s, the physical branch has two critical points and may generate even wave packets propagating in the un-physical direction.

CHAPTER 3. Spline semi-discretizations of the $1-d$ wave equation. In this chapter we analyze and compare the propagation properties for a class of semi-discretizations of the $1-d$ wave equation depending on a parameter $k \in \mathbb{N} \setminus \{0,1\}$ denoting the regularity of the approximated solution, i.e. the numerical solution belongs to $C^k(\mathbb{R})$. Only one type of basis functions supported in intervals of length $(k+2)h$ is used. These methods are called k-refinements (cf. [29]) or cardinal

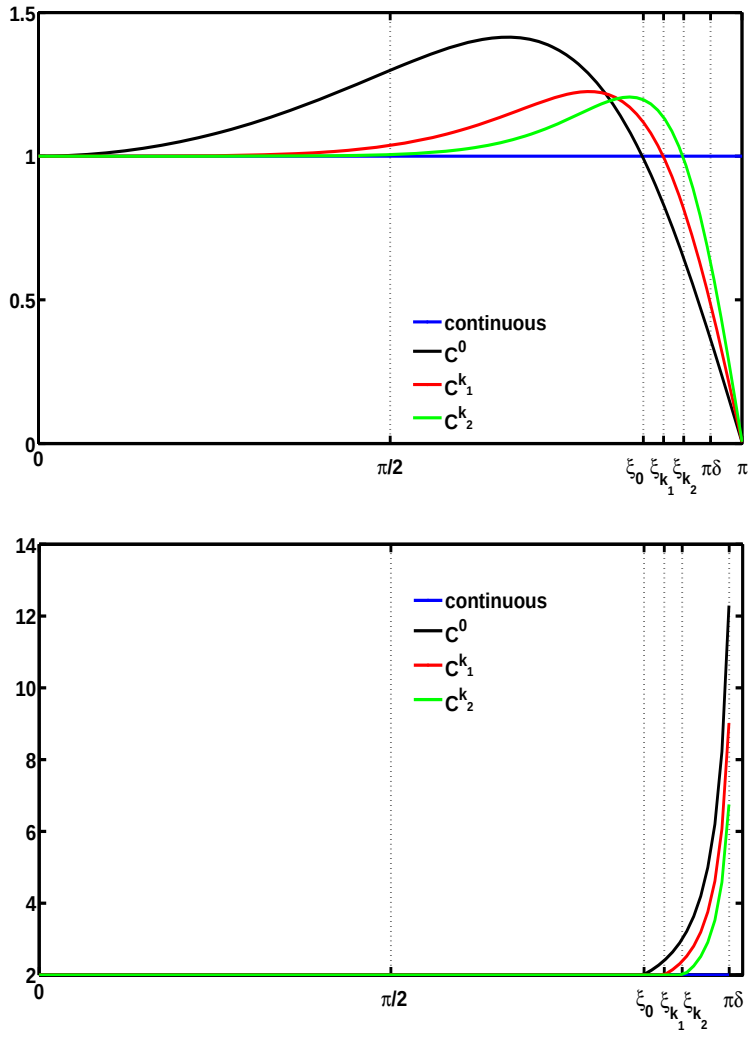

Figure 2.2: Group velocity versus optimal observability time corresponding to the C^k-methods for $k = 0$, $k_1 = 1$ and $k_2 = 2$.

splines ([49], [20]). Despite of their higher regularity and order of approximation to the continuous wave equation compared to the linear classical FEM, these methods maintain a similar pathological behavior. More precisely, for each k, they yield only one dispersion relation with a singular point located at the wave number $\xi = \pi/h$. Thus, pathological high frequency wave packets can be built using the methods of the previous chapters. Without taking filtered initial data, the associated observability constant blows-up at least polynomially at an arbitrary order.

The case $k = 0$ coincides with the P_1-classical finite element scheme and in the limit as $k \to \infty$ one recovers the continuous wave equation. Accordingly, as k increases, the dispersion relation corresponding to a C^k-method adapts better and better to the continuous one that the one corresponding to the P_1-classical finite element method. As one can observe in Figure 2.2, the corresponding group velocity has a maximum point strictly located in $(0, \pi/h)$. It reaches the second time the value one at a wave number $\xi_k \in (\pi/2h, \pi/h)$ and ξ_k increases as function of k. Set T_k^δ to be the optimal observability time derived out of the analysis of the dispersion diagram when the initial data is filtered with parameter $\delta \in (0, 1)$. When a C^k-method is filtered through

a bi-grid algorithm, we observe that the minimal observability time is optimal, $T_k^{1/2} = 2$, for all $k \geq 0$. By the contrary, when the initial data is filtered through a Fourier truncation procedure of parameter $\delta \in (\xi_{k'} h/\pi, 1)$, then $2 < T_{k'}^{\delta} < T_k^{\delta}$, for all $0 \leq k < k'$, which shows that the efficiency of the k-methods increases with k only for the highest frequencies.

CHAPTER 4. Higher-order classical finite element semi-discretizations of the $1-d$ wave equation. This chapter deals with a class of more sophisticated approximations of the $1-d$ wave equation, arising from a space semi-discretization by higher order classical finite element methods on an uniform grid of size $h > 0$. In each computational cell (x_j, x_{j+1}), we add an auxiliary grid, called *serendipity mesh* (cf. [11]), formed by the points $x_j + l\widetilde{h}$, $l = 1, \ldots, k-1$, $\widetilde{h} = h/k$. The P_k classical finite element approximation uses k types of basis functions belonging to $C^0(\mathbb{R})$. One of them represents the nodal points x_j, is supported on (x_{j-1}, x_{j+1}) and takes value one at x_j and zero at each *serendipity point* in $[x_{j-1}, x_{j+1}]$. The remaining $k-1$ basis functions are supported on $[x_j, x_{j+1}]$, take value one at one *serendipity point* and zero at the remaining ones.

For simplicity, we consider only the quadratic approximation for simplicity. The reason for this restriction is that we use explicitly the expressions of the eigenvalues of the Fourier symbol which, for the general case of polynomials of order k, is a $k \times k$ matrix. It is technically complicated to solve the characteristic equation of order k to find the eigenvalues of the symbol depending on the Fourier parameter ξ, and for this reason, we restrict ourselves to the case $k = 2$, for which there exist explicit formulas to find the roots of a polynomial of order k. Nevertheless, a rigorous analysis of this particular case indicates how to deal with the case of a general k.

Despite of the fact that for $C^\infty(\mathbb{R})$ initial data in the continuous model, the performance of P_k-classical finite element approximations increases with respect to k, some singular phenomena appear in what concerns the propagation properties. To be more precise, the P_k semi-discretization of the wave equation yields k dispersion diagrams (an acoustical one and $k-1$ optical ones). On the acoustical one, there exists only a critical point located at $\xi = \pi/h$ and on each optical diagram there exist two singular points at the wave numbers $\xi = 0$ and $\xi = \pi/h$. We have analyzed this kind of behavior for the P_1-SIPG semi-discretization of the wave equation. Thus, by adapting the previous constructions of wave packets propagating arbitrarily slowly, we conclude that the corresponding observability constant blows-up polynomially.

The first filtering mechanism we develop for the P_k-classical finite element method follows the first one we described for the DG methods. Thus, it consists in considering solutions involving only the acoustical branch of frequencies. But one has to eliminate the pathology introduced by the critical point existing on that branch at $\xi = \pi/h$ either by a Fourier truncation or by a bi-grid algorithm.

The filtering mechanism B) we have described for the DG approximation was based on the idea that eliminating the jumps from the numerical initial data should help to remove the pathology generated by the singularity existing on the spurious branch at the wave number $\xi = 0$. The analogous of that for the high-order classical finite element approximation under consideration consists in considering linear numerical initial data. For the quadratic classical finite element approximation, this means that the values of the initial data at a midpoint are averages of their values at the two neighboring nodal points. Once the initial data has been prepared so that most of their energy is concentrated on the acoustical branch, the spurious high frequency components can be filtered out as for previous numerical methods by a Fourier truncation or a bi-grid algorithm.

For $k = 2$, the resulting acoustical dispersion relation approximates better the continuous one. The group velocity corresponding to this physical branch has a maximum point $\xi_k^*(0, \pi/h)$. Denote by $\xi_k \in (\xi_k^*, \pi/h)$ the wave number where the group velocity corresponding to the acoustical branch reaches the value one. We have $\xi_k > \pi/2h$. Let T_k^{δ} be the optimal observability time derived out of the analysis of the acoustical dispersion diagram when the initial data is filtered with parameter $\delta \in (0, 1)$. By considering linear initial data with nodal values given by a bi-grid algorithm, the optimal observability time is $T_k^{1/2} = T = 2$. Nevertheless, by using a Fourier truncation with

$\delta \in (\xi_{k'} h/\pi, 1)$, the corresponding observability time $T^{\delta}_{k'}$ satisfies $2 < T^{\delta}_{k'} < T^{\delta}_{k}$, for all $k < k'$. This means that, from a propagation point of view, the higher order classical finite elements approximations are more efficient at the highest wave numbers.

CHAPTER 5. Further comments and open problems. We close this thesis with a list of Open Problems and future research lines related to the topics we have addressed along the thesis.

Chapter 3

Finite difference approximations of the wave equation

3.1 Introduction

This chapter is aimed to analyze the fine propagation properties of the solutions of finite difference semi-discrete approximations of the wave equation in the whole Euclidean space. In particular, we build high frequency wave packets that propagate according to the group velocity, as predicted in [54]. Our work is primarily motivated by control theoretical issues, but the constructions are of intrinsic interest.

Consider the Cauchy problem associated to the d-dimensional conservative wave equation:

$$\begin{cases} \partial_t^2\phi(x,t) - \Delta\phi(x,t) = 0, & x \in \mathbb{R}^d, t \in (0,T], \\ \phi(x,0) = \phi^0(x), \partial_t\phi(x,0) = \phi^1(x), & x \in \mathbb{R}^d, \end{cases} \tag{3.1}$$

$T > 0$ being a finite time horizon.

The equation (3.1) is well posed in $\dot{H}^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$. More precisely, for each initial data $(\phi^0, \phi^1) \in \dot{H}^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$, as a consequence of the Hille-Yosida Theorem, there exists an unique solution $\phi(x,t)$ of (3.1) in the class $C([0,T], \dot{H}^1(\mathbb{R}^d)) \cap C^1([0,T], L^2(\mathbb{R}^d))$ and the energy of the solution of (3.1),

$$E(\phi^0, \phi^1) = \frac{1}{2}\big(\|\phi(t)\|^2_{\dot{H}^1(\mathbb{R}^d)} + \|\partial_t\phi(t)\|^2_{L^2(\mathbb{R}^d)}\big), \tag{3.2}$$

is conserved in time. Here $\dot{H}^1(\mathbb{R}^d)$ denotes the completion of $\mathcal{D}(\mathbb{R}^d) = C_c^\infty(\mathbb{R}^d)$ with respect to the semi-norm $\|\cdot\|^2_{\dot{H}^1(\mathbb{R}^d)} = \|\nabla\cdot\|^2_{L^2(\mathbb{R}^d)}$.

Given a mesh size $h > 0$, we consider an uniform grid of the whole Euclidean space $x_{\mathbf{j}} = h\mathbf{j}$, $\mathbf{j} \in \mathbb{Z}^d$, and introduce the finite-difference space semi-discretization of wave equation (3.1):

$$\begin{cases} \partial_t^2\phi^h_{\mathbf{j}}(t) - \Delta_h\phi^h_{\mathbf{j}}(t) = 0, \quad \mathbf{j} \in \mathbb{Z}^d, t \in (0,T], \\ \phi^h_{\mathbf{j}}(0) = \phi^{h,0}_{\mathbf{j}}, \quad \partial_t\phi^h_{\mathbf{j}}(0) = \phi^{h,1}_{\mathbf{j}}, \quad \mathbf{j} \in \mathbb{Z}^d, \end{cases} \tag{3.3}$$

where Δ_h is the discrete Laplacian operator defined by

$$\Delta_h f_{\mathbf{j}} = \frac{1}{h^2}\sum_{l=1}^{d}(f_{\mathbf{j}+\mathbf{e}_l} - 2f_{\mathbf{j}} + f_{\mathbf{j}-\mathbf{e}_l}) \tag{3.4}$$

and $(\mathbf{e}_l)_{l=1}^d$ is the canonical basis in $\mathbb{R}^d$.

We will use various discrete versions of the gradient of $\overrightarrow{f}$, defined as $\nabla_h^{\pm}\overrightarrow{f} = (\partial_{h,k}^{\pm}\overrightarrow{f})_{k=1,\cdots,d}$ or $\nabla_h\overrightarrow{f} = (\partial_{h,k}\overrightarrow{f})_{k=1,\cdots,d}$, where

$$\partial_{h,k}^{\pm}\overrightarrow{f} = \pm\frac{\overrightarrow{f}_{\cdot\pm\mathbf{e}_k} - \overrightarrow{f}}{h} \text{ and } \partial_{h,k}\overrightarrow{f} = \frac{\overrightarrow{f}_{\cdot+\mathbf{e}_k} - \overrightarrow{f}_{\cdot-\mathbf{e}_k}}{2h}.$$

We introduce the following two discrete Sobolev spaces:

$$\ell^2(h\mathbb{Z}^d) = \{\overrightarrow{f} \text{ s.t. } \|\overrightarrow{f}\|_{\ell^2(h\mathbb{Z}^d)}^2 := h^d\sum_{\mathbf{j}\in\mathbb{Z}^d}|f_{\mathbf{j}}|^2 < \infty\} \tag{3.5}$$

and

$$\dot{\hbar}^1(h\mathbb{Z}^d) = \{\overrightarrow{f} \text{ s.t. } \|\overrightarrow{f}\|_{\dot{\hbar}^1(h\mathbb{Z}^d)}^2 := \|\nabla_h^+\overrightarrow{f}\|_{\ell^2(h\mathbb{Z}^d)}^2 = \sum_{k=1}^d\|\partial_{h,k}^+\overrightarrow{f}\|_{\ell^2(h\mathbb{Z}^d)}^2 < \infty\}. \tag{3.6}$$

In (3.3), $(\overrightarrow{\phi}^{h,0}, \overrightarrow{\phi}^{h,1}) = (\phi_{\mathbf{j}}^{h,0}, \phi_{\mathbf{j}}^{h,1})_{\mathbf{j}\in\mathbb{Z}^d}$ is usually taken to be a discretization of the continuous initial data $(\phi^0,\phi^1)\in\dot{H}^1(\mathbb{R}^d)\times L^2(\mathbb{R}^d)$. System (3.3) is well-posed in the space $\dot{\hbar}^1(h\mathbb{Z}^d)\times\ell^2(h\mathbb{Z}^d)$.

The semi-discrete total energy associated to the solution $\overrightarrow{\phi}^h(t)$ of the Cauchy problem (3.3),

$$E_h(\overrightarrow{\phi}^{h,0}, \overrightarrow{\phi}^{h,1}) = \frac{1}{2}\big(\|\overrightarrow{\phi}^h(t)\|_{\dot{\hbar}^1(h\mathbb{Z}^d)}^2 + \|\partial_t\overrightarrow{\phi}^h(t)\|_{\ell^2(h\mathbb{Z}^d)}^2\big), \tag{3.7}$$

is a natural discretization of (3.2) and is conserved in time.

The **observability problem** at both continuous and semi-discrete levels consists on determining whether the total energy of solutions can be estimated in terms of the energy concentrated on some subset of the spatial domain where waves propagate.

In the continuous setting (3.1), when the observation region is the complement of a compact set, observability holds. In particular, for all time $T > T^* = 2$, there exists a constant $C(T) > 0$ such that, for all $(\phi^0,\phi^1)\in\dot{H}^1(\mathbb{R}^d)\times L^2(\mathbb{R}^d)$, the following *continuous observability inequality* holds:

$$\boxed{E(\phi^0,\phi^1) \leqslant C(T)\int_0^T\int_{\Omega^d}(|\partial_t\phi(x,t)|^2 + |\nabla_x\phi(x,t)|^2)\,dx\,dt,} \tag{3.8}$$

where $\Omega^d = \mathbb{R}^d\setminus B^d(0,1)$ is the complement of the d-dimensional unit ball. Actually, the observability property (3.8) can be proved for any exterior domain of the form $\Omega^d = \mathbb{R}^d\setminus U^d$, provided $T > T^* := \mathrm{diam}(U^d)$, where $U^d\subset\mathbb{R}^d$ is any bounded open set. In the following, without loss of generality, we will focus only on the case $U^d = B^d(0,1)$.

The observability problem is motivated by controllability issues. More precisely, by means of the Hilbert Uniqueness Method (HUM) (see [36]), (3.8) is equivalent to the fact that for all $T > T^*$ and initial data $(u^0,u^1)\in\dot{H}^1(\mathbb{R}^d)\times L^2(\mathbb{R}^d)$, there exists a control function $f\in L^2(\Omega^d\times(0,T))$ such that the solution of the inhomogeneous Cauchy problem

$$\begin{cases} \partial_t^2u(x,t) - \Delta u(x,t) = f(x,t)\chi_{\Omega^d}(x), & x\in\mathbb{R}^d,\ t\in(0,T],\\ u(x,0) = u^0(x),\ \partial_t u(x,0) = u^1(x),\ x\in\mathbb{R}^d \end{cases} \tag{3.9}$$

satisfies

$$u(x,T) = \partial_t u(x,T) = 0,\ \forall x\in\mathbb{R}^d. \tag{3.10}$$

Here χ_{Ω^d} denotes the characteristic function of the set Ω^d.

These issues are by now well understood for the continuous wave equation and have been the object of an intensive research. We refer to [60] for a recent survey on this and closely related issues. In particular, it is well known that observability holds under the **sharp Geometric**

Control Condition (GCC) (see [7]). This condition requires that *all non-diffractive rays of Geometric Optics enter the observation set during the observability time.*

To be more precise, let us define in a rigorous manner the notion of *rays of Geometric Optics*. If $H(x,t,\xi,\tau)=\tau^2-p(x,t,\xi,\tau)$ is the Hamiltonian associated to some continuous (or even semi-discrete) second-order hyperbolic problem, $p(x,t,\xi,\tau)$ being the principal symbol of an elliptic operator (obtained by applying the continuous Fourier transform on the spatial operator when one deals with a continuous operator and the semi-discrete Fourier transform when one deals with a discrete operator on an uniform grid), then the *bi-characteristic rays* are parametrized curves $(x(s),t(s),\xi(s),\tau(s))$ verifying the following Hamiltonian system:

$$\left\{\begin{array}{l} \dot{x}(s)=\nabla_\xi H(x(s),t(s),\xi(s),\tau(s)), \quad \dot{t}(s)=\partial_\tau H(x(s),t(s),\xi(s),\tau(s)),\\ \dot{\xi}(s)=-\nabla_x H(x(s),t(s),\xi(s),\tau(s)), \quad \dot{\tau}(s)=-\partial_t H(x(s),t(s),\xi(s),\tau(s)),\\ (x(0),t(0),\xi(0),\tau(0))=(x^0,t^0,\xi^0,\tau^0), \end{array}\right. \tag{3.11}$$

starting from initial data in the characteristic manifold $H(x^0,t^0,\xi^0,\tau^0)=0$. The rays of Geometric Optics are projections of the bi-characteristic rays on the (x,t)-plane. For the classical Laplacian, $-\Delta$, the Hamiltonian is $H(x,t,\xi,\tau)=\tau^2-|\xi|^2$ and, consequently, the rays of Geometric Optics are straight lines $x(t)=x^0\pm\kappa t$, with κ a vector s.t. $|\kappa|=1$. This means that the velocity of propagation of rays in the continuous case is always one. This explains why the time $T^*=2$ is the minimal one for the observability inequality (3.8) to hold, i.e. it is the time needed for the GCC property to hold when the observation set is the exterior of the unit ball.

When the GCC is not satisfied, the property of observability fails because of the existence of Gaussian beam solutions localized around a bi-characteristic ray that escapes the observation region during the time interval $[0,T]$ (see, for instance, [35],[37], [38], [46], [48] and [56]). One of the goals of this chapter is to extend the Gaussian beam construction, known by now in the context of continuous wave equations, to the semi-discrete framework.

The analysis in this chapter is motivated by the analogue of (3.8) for classical numerical approximation schemes like the finite difference one and its control theoretical consequences. In a first approximation to the topic, one could think that numerical schemes that converge in the classical sense of numerical analysis should share with the continuous wave equation the observability inequality (3.8) and this uniformly on the mesh-size parameter. But it is by now well known that (3.8) fails to be uniform with respect to the mesh size under some numerical approximations of PDEs because of the pathological behavior of spurious high frequency numerical solutions. In those cases, the observability constant blows-up when the mesh size tends to zero. This occurs, for instance, for the classical finite difference and finite element discretization schemes on particular bounded domains like d-dimensional cubes. We refer to [59] for a recent survey on this topic where several examples of blow-up of the observability constant are shown in terms of the Fourier representation of solutions. Similar pathological high frequency behaviors have also been observed in other contexts as, for instance, in what concerns the Strichartz dispersive estimates for finite difference approximation schemes of the Schrödinger equation (cf. [31]). Some possible remedies for avoiding the high frequency spurious oscillations have also been developed: Tychonoff regularization, multi-grid methods, mixed finite element methods ([17],[18]), numerical viscosity (see [43]) and Fourier filtering of high frequencies.

This chapter is devoted to analyze in more detail the blow-up rate of the observability constant in the observability inequality associated to the semi-discrete problem (3.3), by building concentrated high frequency wave packets. To make the problem more precise, we fix a finite time $T>0$ and consider the best constant $C_h(T)>0$ such that for all $(\overrightarrow{\phi}^{h,0},\overrightarrow{\phi}^{h,1})\in \dot{h}^1(h\mathbb{Z}^d)\times\ell^2(h\mathbb{Z}^d)$ the following *discrete observability inequality* holds for the corresponding solution of (3.3):

$$\boxed{E_h(\overrightarrow{\phi}^{h,0},\overrightarrow{\phi}^{h,1})\leq C_h(T)h^d\sum_{x_{\mathbf{j}}\in\Omega^d}\int_0^T(|\partial_t\phi_{\mathbf{j}}(t)|^2+|\nabla_h^+\phi_{\mathbf{j}}(t)|^2)\,dt.} \tag{3.12}$$

It is easy to see that the constant $C_h(T)$ is well defined and finite for all $T, h > 0$. Indeed, taking into account that the number of nodes of the mesh that escape the zone of observation Ω^d is finite, the inequality (3.12) is simply a consequence of the following unique continuation property: If ϕ, the solution of (3.3) is such that $\partial_t \phi_{\mathbf{j}}(t) = 0$, for all $\mathbf{j} \in \{\mathbf{j} : x_{\mathbf{j}} \in \Omega^d\}$ and $0 < t < T$, then, necessarily $\phi \equiv 0$. This unique continuation property is easy to prove by a classical argument of propagation of the zero level set of solutions of the semi-discrete scheme (we refer to [19], where this is done in the context of the Laplace operator).

But, as mentioned above, for all $T > 0$, the observability constant $C_h(T)$ blows-up as $h \to 0$. Our constructions of highly concentrated wave packets allow obtaining lower bounds on the blow-up rate of $C_h(T)$. These pathological solutions consist on wave packets corresponding to initial data obtained by modulating high frequency plane waves by functions concentrated in the physical or in the Fourier space. As mentioned above, to some extent, this chapter makes rigorous the argument leading to the introduction of the notion of group velocity in [15], [55], [12], [54], describing the propagation of wave packets, i.e. of short wavelength oscillations modulating a slower varying envelope, within a dispersive medium.

In order to give a precise mathematical definition of the group velocity, we recall here, for the sake of completeness, the definition and the main properties of the semi-discrete Fourier transform (SDFT) at the scale h, which is one of the main tools used to carefully analyze the semi-discrete system (3.3). We refer to [53] and [51] for further details.

Define $\Pi_h^d := [-\pi/h, \pi/h]^d$. For $x_{\mathbf{j}} = \mathbf{j}h$, the SDFT at the scale h of a sequence $\overrightarrow{f}$ is a Π_h^d-periodic function on $\mathbb{R}^d$, given by

$$\widehat{f}^h(\xi) = h^d \sum_{\mathbf{j} \in \mathbb{Z}^d} f_{\mathbf{j}} \exp(-i\xi \cdot x_{\mathbf{j}}), \text{ for all } \xi \in \Pi_h^d. \tag{3.13}$$

The SDFT at the scale h is an isometry between $\ell^2(h\mathbb{Z}^d)$ and $L^2(\Pi_h^d)$ as one can easily observe from the discrete Parseval identity:

$$\|\overrightarrow{f}\|^2_{\ell^2(h\mathbb{Z}^d)} = \frac{1}{(2\pi)^d} \int\limits_{\Pi_h^d} |\widehat{f}^h(\xi)|^2 \, d\xi. \tag{3.14}$$

The sequence $\overrightarrow{f}$ can be recovered from its SDFT using the inversion formula:

$$f_{\mathbf{j}} = \frac{1}{(2\pi)^d} \int\limits_{\Pi_h^d} \widehat{f}^h(\xi) \exp(i\xi \cdot x_{\mathbf{j}}) \, d\xi. \tag{3.15}$$

By applying the SDFT at scale h, system (3.3) can be transformed into the following second-order ODE depending on the Fourier parameter ξ:

$$\begin{cases} \partial_t^2 \widehat{\phi}^h(\xi, t) + \omega_{d,h}^2(\xi) \widehat{\phi}^h(\xi, t) = 0, & \xi \in \Pi_h^d, t \in (0, T], \\ \widehat{\phi}^h(\xi, 0) = \widehat{\phi}^{h,0}(\xi), \quad \partial_t \widehat{\phi}^h(\xi, 0) = \widehat{\phi}^{h,1}(\xi), & \xi \in \Pi_h^d, \end{cases} \tag{3.16}$$

whose solution is given by

$$\widehat{\phi}^h(\xi, t) = \sum_{\pm} \frac{1}{2} \Big(\widehat{\phi}^{h,0}(\xi) \pm \frac{\widehat{\phi}^{h,1}(\xi)}{i\omega_{d,h}(\xi)} \Big) \exp(\pm it\omega_{d,h}(\xi)), \tag{3.17}$$

where

$$\omega_{d,h}^2(\xi) = \sum_{k=1}^d \omega_h^2(\xi_k), \quad \text{with} \quad \omega_h(\xi_k) = \frac{2}{h} \sin\big(\frac{\xi_k h}{2}\big), \quad \xi \in \Pi_h^d. \tag{3.18}$$

In (3.18), $\omega_{d,h}(\xi)$ is the multi-dimensional dispersion relation associated to the numerical scheme (3.3), while $\omega_h(\xi_k)$ is the one-dimensional one. Observe that, by setting $\eta = h\xi \in \Pi_1^d$, we have $\omega_{d,1}(\eta) = h\omega_{d,h}(\xi)$.

As $h \to 0$, for a fixed ξ, $\omega_{d,h}(\xi)$ converges to the dispersion relation $|\xi|$ of the continuous wave equation. This is in agreement with the consistence of the finite difference scheme in the classical sense of numerical analysis of PDEs.

Particularizing system (3.11) to the Hamiltonian $H_h(x,t,\xi,\tau) = \tau^2 - \omega_{d,h}^2(\xi)$, we obtain that the semi-discrete rays of Geometric Optics corresponding to the discrete problem (3.3) are of the form

$$x_h^{\pm}(t) = x^* \pm \nabla\omega_{d,h}(\xi)t, \quad \xi \in \Pi_h^d, \tag{3.19}$$

where

$$\nabla\omega_{d,h}(\xi) = \frac{1}{h\omega_{d,h}(\xi)}(\sin(\xi_1 h), \cdots, \sin(\xi_d h)).$$

This is precisely the group velocity corresponding to the approximation (3.3) of the wave equation. Observe that in our particular case, the velocity of propagation $|\nabla\omega_{d,h}(\xi)|$ vanishes for all $\xi \in \{\pm\pi/h, 0\}^d \setminus \{0\}$ (see Figures 3.1 and 3.2 for graphical representations of the velocities of propagation $|\nabla\omega_{d,1}(\eta)|$ in the one and two-dimensional cases), a behavior which does not happen in the continuous case. Formally, the fact that the velocity of propagation has not a strictly positive lower bound for all range of frequencies is the obstruction for the observability inequality (3.12) to be uniform as $h \to 0$.

In this paper we make rigorous constructions of wave packets traveling with pathological group velocities arbitrarily close to zero which lead to the divergence of the observability constant and, moreover to lower bounds on the divergence rate.

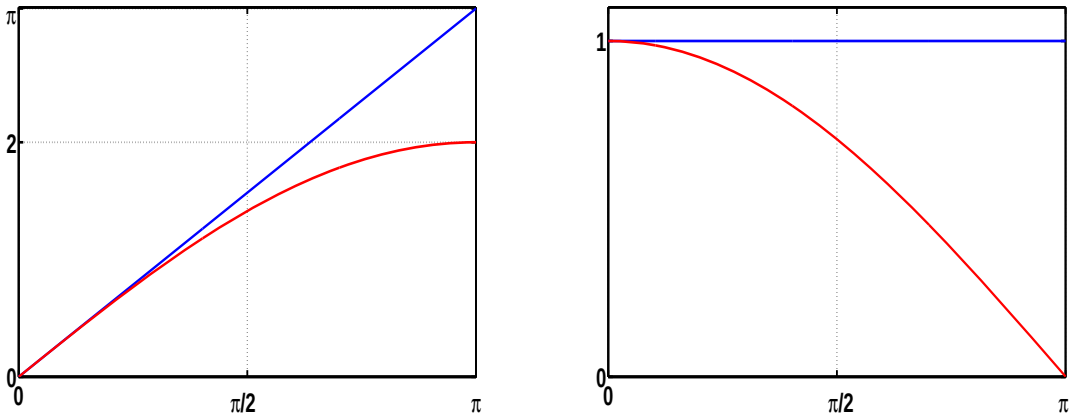

Figure 3.1: Dispersion relation versus group velocity for the $1-d$ continuous wave equation (blue) and for its finite difference semi-discretization (red), with $h = 1$.

Using the Fourier representation of the discrete gradient,

$$\partial_{h,k}^{\pm} f_{\mathbf{j}} = \frac{1}{(2\pi)^d} \int_{\Pi_h^d} \widehat{f}^h(\xi) \exp(i\xi \cdot x_{\mathbf{j}}) \exp\Big(\pm \frac{i\xi_k h}{2}\Big) i\omega_h(\xi_k)\, d\xi, \tag{3.20}$$

and the Parseval identity (3.14), we find that the semi-discrete energy can be represented as

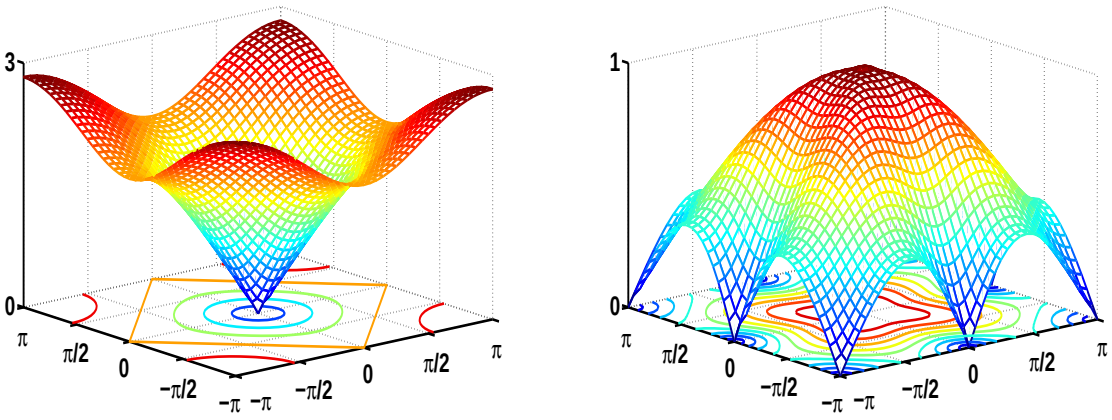

Figure 3.2: Dispersion relation versus velocity of propagation for the $2-d$ finite-difference semi-discretization of the wave equation for $h=1$.

follows:

$$E_h(\overrightarrow{\phi}^{h,0},\overrightarrow{\phi}^{h,1})=\frac{1}{2(2\pi)^d}\int\limits_{\Pi_h^d}\left(|\widehat{\phi^h}(\xi,t)|^2\omega_{d,h}^2(\xi)+|\partial_t\widehat{\phi^h}(\xi,t)|^2\right)d\xi. \tag{3.21}$$

In the context of the finite difference semi-discretization of the $1-d$ wave equation on a finite interval, it was proved that the corresponding observability constant blows-up exponentially as $h\to 0$ for all finite $T>0$ (see [42]). The proof was based on Fourier series representation of solutions and fine estimates on bi-orthogonal sequences.

The results we present in this chapter are slightly weaker in the sense that we can only prove the divergence of the observability constant at an arbitrarily large polynomial rate. But the constructions we develop have several advantages with respect to previously existing ones. In particular, they are based on the analysis of the symbol of the operator and do not require an explicit knowledge of the spectrum of the system. Therefore, they can be applied both to the Cauchy problem on the real line and to boundary-value problems with various boundary conditions (cf. [24]).

We now describe the content of the chapter and the main results we obtain:

In Section 3.2, we prove that, by using the results of Micu in [42], at least in the one-dimensional case, the observability constant $C_h(T)$ in (3.12) blows-up exponentially. Nevertheless, we have no information on the solutions $\overrightarrow{\phi}^h(t)$ of (3.3) for which the exponential blow-up holds (concerning, for instance, their shape). Also, we do not know if this exponential blow-up is maintained for the multi-dimensional version, which motivates our analysis along this chapter.

In Section 3.3, we construct exact solutions of the semi-discrete equation (3.3) concentrated along a ray of Geometric Optics. There are two main ingredients of this construction:

i. A wave number $\xi_0=\eta_0/h\in\Pi_h^d\setminus\{0\}$ around which the SDFT of the numerical solution is concentrated. It determines the ray. More precisely, the choice of the wave number and of the observability time T is such that the ray of slope $\nabla\omega_{d,h}(\xi_0)=\nabla\omega_{d,1}(\eta_0)$ does not enter the observation region in time T.

ii. A smooth function $\widehat{\sigma}\in C_c^\infty(B^d(0,1))$. The initial data is of type $\gamma^{-d/2}\widehat{\sigma}(\gamma^{-1}(\xi-\xi_0))$, with $\gamma=\gamma(h)$ such that $\lim\limits_{h\to 0}\gamma(h)h=0$ and $\lim\limits_{h\to 0}\gamma(h)=+\infty$.

In this way, the initial data we construct are high frequency wave packets at concentrated the wave number $\xi=\xi_0=\eta_0/h$ with a width of order γ. This limits drastically the spread of

the wave packet and allows to concentrate its energy around the chosen ray and to reduce the added dispersive effects that the non-trivial Hessian matrix of the dispersion relation introduces. Simultaneously, by the uncertainty principle, this forces the wave packet, in the physical space, to have a spread factor of order γ^{-1}, with $1 >> \gamma^{-1} >> h$ as $h \to 0$, which is asymptotically larger than the characteristic size h of the mesh. This is natural, in the sense that the numerical effects are only detected when an infinite number of nodes enter asymptotically in the determination of the data of the solution.

In Section 3.4, some qualitative properties of the solutions we construct are analyzed: in particular, the dispersive properties and their asymmetric time-evolution.

The analysis in Section 3.5 reveals the existence of a scale $\gamma = h^{-1/2}$ in the case when an analytic dispersion relation is approximated by a quadratic one. The general case when the complete Taylor expansion of the dispersion relation is considered seems make appearing all the scales $\gamma = h^{-k/(k+1)}$, $\forall k \in \mathbb{N}$. The techniques involved in the proof of the results in this section are of asymptotic type, known as WKB expansions.

In Section 3.6, we extend the Gaussian beams construction existing for continuous hyperbolic equuations ([46], [38]) to the corresponding finite difference semi-discretization. This type of ansatz provides an asymptotic expansion of the numerical solution. For this reason, it needs a correction argument to obtain the behavior of the observability constant associated to the exact solution of the semi-discrete wave equation with the same initial data as in the asymptotic expansion. The observability constant associated to the Gaussian beam blows-up exponentially, but the one corresponding to the correction blows-up polynomially at any order.

3.2 Exponentially concentrated $1-d$ numerical waves

In this section, based on the results of Micu in [42], we show the existence of exponentially concentrated solutions of (3.3) in the one-dimensional case. Through this section, $d = 1$. More precisely, we prove:

Proposition 3.2.1. *Consider $N \in \mathbb{N}$ and $h = 1/(N+1)$. For all finite time $T > 0$, there exists a solution of problem (3.3), $\overrightarrow{\phi}^h(t)$, such that the associated observability constant $C_h(T)$ defined by (3.12) satisfies $C_h(T) \geq ch\exp(\sqrt{N})/T^2$, where $c > 0$ does not depend on h.*

Proof. Set $\Lambda_N := [-N, N] \cap \mathbb{Z}$ and the spaces

$$\ell^2 = \{\overrightarrow{f} = (f_j)_{|j|\leq N} : \|\overrightarrow{f}\|^2_{\ell^2} = h\sum_{j\in\Lambda_N} |f_j|^2 < \infty\}$$

and

$$\hbar_0^1 = \{\overrightarrow{f} = (f_j)_{|j|\leq N+1} : f_{\pm(N+1)} = 0, \|\overrightarrow{f}\|^2_{\hbar_0^1} := \|\partial_h^+\overrightarrow{f}\|^2_{\ell^2} < \infty\}.$$

Consider the following semi-discrete problem

$$\begin{cases} \partial_t^2 u_j(t) - \partial_h^2 u_j(t) = 0, & j \in \Lambda_N, t \in (0,T) \\ u_{\pm(N+1)}(t) = v^h_{\pm 1}(t), & t \in (0,T), \\ u_j(0) = u_j^0,\ \partial_t u_j(0) = u_j^1, & j \in \Lambda_N. \end{cases} \tag{3.22}$$

The exact controllability problem from the boundary requires the existence of a time $T^* > 0$ such that for all $T > T^*$ and $(\overrightarrow{u}^0, \overrightarrow{u}^1) \in (\hbar_0^1 \times \ell^2)'$, one is able to find controls $v^h_{\pm 1} \in L^2(0,T)$ such that the solution of (3.22) verifies $u_j(T) = \partial_t u_j(T) = 0$, $\forall j \in \Lambda_N$. In [42], it was proved that there exist controls $v^h_{\pm 1}(t)$ s.t. $\sum\limits_{\pm} \|v^h_{\pm 1}\|_{L^2(0,T)} \geq \exp(\sqrt{N})$. One of such controls is the last element of any bi-orthogonal sequence to the family of complex exponentials $\exp(it\omega_h(k\pi))_{k\in\Lambda_N}$. By HUM,

the exact controllability problem (3.22) is equivalent to an observability problem for the adjoint problem to (3.22),

$$\begin{cases} \partial_t^2\phi_j(t) - \partial_h^2\phi_j(t) = 0, & j \in \Lambda_N, t \in (0,T) \\ \phi_{\pm(N+1)}(t) = 0, & t \in (0,T), \\ \phi_j(0) = \phi_j^0,\ \partial_t\phi_j(0) = \phi_j^1, & j \in \Lambda_N. \end{cases} \tag{3.23}$$

As expected, the total energy corresponding to the solution of (3.23),

$$E_h(\overrightarrow{\phi}^{h,0}, \overrightarrow{\phi}^{h,1}) = \frac{1}{2}\big(\|\overrightarrow{\phi}^h(t)\|^2_{\hbar_0^1} + \|\overrightarrow{\phi}^h(t)\|^2_{\ell^2}\big),$$

is constant in time. The observability constant is defined as

$$C_h(T) = \sup_{(\overrightarrow{\phi}^0, \overrightarrow{\phi}^1)\in\hbar_0^1\times\ell^2} \frac{E_h(\overrightarrow{\phi}^{h,0}, \overrightarrow{\phi}^{h,1})}{\int_0^T \Big(\Big|\frac{\phi_N}{h}\Big|^2 + \Big|\frac{\phi_N}{h}\Big|^2\Big)\,dt}. \tag{3.24}$$

Taking into account that, by HUM, $C_h(T) = \sup_{v_\pm^h \text{ admissible}} \sum_{\pm} \|v_{\pm 1}^h\|_{L^2(0,T)}$, we conclude that there exists a solution $(\phi^{exp}(t))_{j\in\lambda_N}$ of the adjoint problem (3.23) for which the observability constant blows-up at least at order $\exp(\sqrt{N})$.

Let us denote by $\phi_j^{ext}(t)$ the extension by zero for indices $|j| > N+1$ of $\phi_j^{exp}(t)$. The sequence $(\phi_j^{ext}(t))_{j\in\mathbb{Z}}$ verifies the non-homogeneous semi-discrete wave equation on the whole real line

$$\begin{cases} \partial_t^2\phi_j^{ext}(t) - \partial_h^2\phi_j^{ext}(t) = f_j(t), & j \in \mathbb{Z}, t \in (0,T) \\ \phi_j^{ext}(0) = \phi_j^{ext,0},\ \partial_t\phi_j^{ext}(0) = \phi_j^{ext,1}, & j \in \mathbb{Z}, \end{cases} \tag{3.25}$$

where $f_j(t) = 0$ for $|j| \neq N+1$ and $f_{\pm(N+1)}(t) = -\phi_{\pm N}/h^2$. Consider $(\phi_j^{corr}(t))_{j\in\mathbb{Z}}$ to be the solution of the following corrector problem

$$\begin{cases} \partial_t^2\phi_j^{corr}(t) - \partial_h^2\phi_j^{corr}(t) = -f_j(t), & j \in \mathbb{Z}, t \in (0,T) \\ \phi_j^{corr}(0) = \partial_t\phi_j^{corr}(0) = 0, & j \in \mathbb{Z}. \end{cases} \tag{3.26}$$

Then $\phi_j(t) := \phi_j^{ext}(t) + \phi_j^{corr}(t)$ is the solution of (3.3) corresponding to the initial data $(\overrightarrow{\phi}^{ext,0}, \overrightarrow{\phi}^{ext,1})$. We use the following lemma:

Lemma 3.2.1. *For all finite $T > 0$, the following estimate holds for the energy of $\overrightarrow{\phi}^{corr}$:*

$$\max_{t\in[0,T]} E_h(\overrightarrow{\phi}^{corr}(t), \partial_t\overrightarrow{\phi}^{corr}(t)) \le 2Th^{-1}\int_0^T \Big(\Big|\frac{\phi_N^{exp}(t)}{h}\Big|^2 + \Big|\frac{\phi_{-N}^{exp}(t)}{h}\Big|^2\Big)\,dt.$$

Taking into account that $\phi_j^{ext}(t) = 0$ for all $|x_j| > 1$, therefore $\partial_t\phi_j^{ext}(t) = 0$, and that $E_h(\overrightarrow{\phi}^{exp,0}, \overrightarrow{\phi}^{exp,1}) = E_h(\overrightarrow{\phi}^{ext,0}, \overrightarrow{\phi}^{ext,1})$, by Lemma 3.2.1, we have

$$\begin{aligned} \frac{E_h(\overrightarrow{\phi}^{h,0}, \overrightarrow{\phi}^{h,1})}{h\sum_{|x_j|>1}\int_0^T |\partial_t\phi_j(t)|^2} &\ge \frac{E_h(\overrightarrow{\phi}^{ext,0}, \overrightarrow{\phi}^{ext,1})}{2h\sum_{|x_j|>1}\int_0^T \big(|\partial_t\phi_j^{ext}(t)|^2 + |\partial_t\phi_j^{corr}(t)|^2\big)} = \frac{E_h(\overrightarrow{\phi}^{exp,0}, \overrightarrow{\phi}^{exp,1})}{4\int_0^T E_h(\overrightarrow{\phi}^{corr}(t), \partial_t\overrightarrow{\phi}^{corr}(t))\,dt} \\ &\ge \frac{h}{8T^2}\frac{E_h(\overrightarrow{\phi}^{exp,0}, \overrightarrow{\phi}^{exp,1})}{\int_0^T \Big(\Big|\frac{\phi_N^{exp}(t)}{h}\Big|^2 + \Big|\frac{\phi_{-N}^{exp}(t)}{h}\Big|^2\Big)\,dt.} \ge \frac{h\exp(\sqrt{N})}{8T^2}, \end{aligned}$$

which concludes the proof of Proposition 3.2.1. □

Proof. Let us multiply (3.26) by $h\partial_t \phi_j^{corr}(t)$, summ in $j \in \mathbb{Z}$ and integrate in time. Then by Cauchy-Schwartz inequality we have:

$$E_h(\overrightarrow{\phi}^{corr}(t), \partial_t \overrightarrow{\phi}^{corr}(t)) \leq \left(h \sum_{j\in\mathbb{Z}} \int_0^t |f_j(s)|^2 \, ds \right)^{1/2} \left(h \sum_{j\in\mathbb{Z}} \int_0^t |\partial_t \phi_j^{corr}(s)|^2 \, ds \right)^{1/2}$$
$$\leq \left(h \sum_{j\in\mathbb{Z}} \int_0^T |f_j(t)|^2 \, dt \right)^{1/2} \left(2T \max_{t\in[0,T]} E_h(\overrightarrow{\phi}^{corr}(t), \partial_t \overrightarrow{\phi}^{corr}(t)) \right)^{1/2}.$$

The right hand side in the above inequality does not depend on t. Then, by taking maximum in the left hand side and dividing by the square root of the energy, we obtain

$$\max_{t\in[0,T]} E_h(\overrightarrow{\phi}^{corr}(t), \partial_t \overrightarrow{\phi}^{corr}(t)) \leq 2Th \sum_{j\in\mathbb{Z}} \int_0^T |f_j(t)|^2 \, dt = 2Th^{-1} \int_0^T \left(\left| \frac{\phi_N^{exp}(t)}{h} \right|^2 + \left| \frac{\phi_{-N}^{exp}(t)}{h} \right|^2 \right) dt.$$

□

3.3 Example of polynomially concentrated wave packets

3.3.1 Main results

Theorem 3.3.1. *Fix $T > 0$ and consider a wave number $\eta_0 = h\xi_0 \in \Pi_1^d \setminus \{0\}$ and a point $x^* \in B^d(0,1)$ such that the semi-discrete GCC is not verified, i.e.*

$$|x_h(t)| = |x^* - t\nabla_\eta \omega_{d,1}(\eta_0)| < 1, \quad \forall t \in [0,T]. \tag{3.27}$$

Consider $\widehat{\sigma} \in C_c^\infty(B^d(0,1))$ and $\gamma = \gamma(h) > 0$ to be a function such that the following two condition are verified

$$\lim_{h\to 0} \gamma(h) = +\infty \text{ and } \lim_{h\to 0} h\gamma(h) = 0. \tag{3.28}$$

In the semi-discrete wave equation (3.3), the initial data $(\overrightarrow{\phi}^{h,0}, \overrightarrow{\phi}^{h,1})$ are such that their SDFTs verify the following two requirements:

$$\widehat{\phi}^{h,0}(\xi) = \gamma^{-d/2} \widehat{\sigma}(\gamma^{-1}(\xi - \xi_0)) \frac{1}{i\omega_{d,h}(\xi)} \exp(-i\xi \cdot x^*) \text{ and } \widehat{\phi}^{h,1}(\xi) = i\omega_{d,h}(\xi)\widehat{\phi}^{h,0}(\xi). \tag{3.29}$$

Then for all $\alpha > 0$, there exists a constant $c_\alpha = c_\alpha(T, \widehat{\sigma}, \eta_0) > 0$ (not depending on h) such that the discrete observability constant $C_h(T)$ in (3.12) satisfies

$$C_h(T) \geq c_\alpha \gamma^\alpha.$$

Proof. **The total energy is constant.** Indeed, due to the two requirements on the initial data (3.29), the total energy simplifies with respect to the more general one introduced in (3.21) as follows:

$$E_h(\overrightarrow{\phi}^{h,0}, \overrightarrow{\phi}^{h,1}) = \frac{1}{(2\pi)^d} \int_{\Pi_h^d} |\widehat{\phi}^{h,0}(\xi)|^2 \omega_{d,h}^2(\xi) \, d\xi = \frac{1}{(2\pi)^d} \int_{B^d(\xi_0,\gamma)} \gamma^{-d} |\widehat{\sigma}(\gamma^{-1}(\xi - \xi_0))|^2 \, d\xi$$
$$= \frac{1}{(2\pi)^d} \int_{B^d(0,1)} |\widehat{\sigma}(\eta)|^2 \, d\eta = \text{constant}.$$

Then, without loss of generality, we assume that the total energy is one.

In what follows, we analyze the energy concentrated in the set

$$\Omega_\delta(t) := \{x \in \mathbb{R}^d : |x - x_h(t)| > \delta\},$$

for some fixed $\delta > 0$ and all $t \in [0, T]$. More precisely, we prove that:

Lemma 3.3.1. *Under the hypothesis of Theorem 3.3.1, for all fixed $\delta > 0$ and all $N \in \mathbb{N}$, $N > d/2$, there exists a constant $C_N(T, \widehat{\sigma}, \delta, \eta_0) > 0$ independent on h, such that for all $t \in [0, T]$, the following estimate holds:*

$$h^d \sum_{x_{\mathbf{j}} \in \Omega_\delta(t)} (|\partial_t \phi_{\mathbf{j}}(t)|^2 + |\nabla_h^+ \phi_{\mathbf{j}}(t)|^2) \leq C_N(T, \widehat{\sigma}, \delta, \eta_0) \gamma^{-(2N-d)}. \tag{3.30}$$

The conclusion of Theorem 3.3.1 holds since T and ξ_0 are such that the discrete GCC (3.27) is not verified. Since the GCC does not hold, there exists $\delta > 0$ such that the whole cylinder $\{x : |x - x_h(t)| < \delta\} \times [0, T]$ is contained in $B^d(0, 1) \times [0, T]$. The following inequality is obvious:

$$h^d \int_0^T \sum_{x_{\mathbf{j}} \in \Omega} (|\partial_t \phi_{\mathbf{j}}(t)|^2 + |\nabla_h^+ \phi_{\mathbf{j}}(t)|^2)\, dt \leq h^d \int_0^T \sum_{x_{\mathbf{j}} \in \Omega_\delta(t)} (|\partial_t \phi_{\mathbf{j}}(t)|^2 + |\nabla_h^+ \phi_{\mathbf{j}}(t)|^2)\, dt.$$

□

Before proceeding to the proof of Lemma 3.3.1, let us define the function

$$\phi(x, t) = \frac{1}{(2\pi)^d} \int_{B^d(\xi_0, \gamma)} \gamma^{-d/2} \widehat{\sigma}(\gamma^{-1}(\xi - \xi_0)) \frac{1}{i\omega_{d,h}(\xi)} \exp(it\omega_{d,h}(\xi) + i\xi \cdot (x - x^*))\, d\xi \tag{3.31}$$

and observe that this function is an interpolation of the solution of (3.3) with the initial data (3.29) in the sense that $\phi_{\mathbf{j}}(t) = \phi(x_{\mathbf{j}}, t)$. Also, let us define the following phase function

$$\psi(\eta, x, t) = t\omega_{d,1}(\eta + \eta_0) + x \cdot (\eta + \eta_0). \tag{3.32}$$

By the change of variable $\xi - \xi_0 = \gamma\eta$, the function ϕ can be written as

$$\phi(x, t) = \frac{\gamma^{d/2}}{(2\pi)^d} \int_{B^d(0,1)} \widehat{\sigma}(\eta) \exp\Big(\frac{i}{h}\psi(\gamma h\eta, x - x^*, t)\Big) \frac{h}{i\omega_{d,1}(\eta_0 + \eta\gamma h)}\, d\eta. \tag{3.33}$$

Then the time derivative and the discrete space gradient of ϕ are given explicitly by

$$\partial_t \phi(x, t) = \frac{\gamma^{d/2}}{(2\pi)^d} \int_{B^d(0,1)} \widehat{\sigma}(\eta) \exp\Big(\frac{i}{h}\psi(\gamma h\eta, x - x^*, t)\Big)\, d\eta. \tag{3.34}$$

and

$$\nabla_h^+ \phi(x, t) = \frac{\gamma^{d/2}}{(2\pi)^d} \int_{B^d(0,1)} \widehat{\sigma}(\eta) \Big(\frac{\omega_1(\eta_{0,k} + \gamma h\eta_k)}{\omega_{d,1}(\eta_0 + \eta\gamma h)} \exp\Big(\frac{i}{2}(\eta_{0,k} + \gamma h\eta_k)\Big)\Big)_{1 \leq k \leq d} \exp\Big(\frac{i}{h}\psi(\gamma h\eta, x - x^*, t)\Big)\, d\eta. \tag{3.35}$$

The following result allows us to replace discrete sums by integrals:

Lemma 3.3.2. *For all $f \in W^{1,1}(\mathbb{R}^d) \cap C(\mathbb{R}^d)$ and all open set $O \subset \mathbb{R}^d$, the following estimate holds:*

$$\Big| h^d \sum_{x_{\mathbf{j}} \in O} f(x_{\mathbf{j}}) - \int\limits_{\bigcup\limits_{x_{\mathbf{j}} \in O} C_{\mathbf{j}}} f(x)\, dx \Big| \leq c(d) h \int\limits_{O + B^d(0, h\sqrt{d}/2)} |\nabla f(x)|\, dx, \tag{3.36}$$

where $C_{\mathbf{j}}$ is the d-dimensional cube centered in $x_{\mathbf{j}}$ with edge length h.

Proof of Lemma 3.3.1. Using Lemma 3.3.2 and the Cauchy-Schwartz inequality, we obtain

$$h^d \sum_{x_{\mathbf{j}} \in \Omega_\delta(t)} (|\partial_t \phi(x_{\mathbf{j}}, t)|^2 + |\nabla_h^+ \phi(x_{\mathbf{j}}, t)|^2) \leq I_1(t) + 2c(d) h (I_1(t))^{1/2} (I_2(t))^{1/2}, \tag{3.37}$$

with

$$I_1(t) = \int\limits_{\Omega_{\delta - h\sqrt{d}/2}} (|\partial_t \phi(x,t)|^2 + |\nabla_h^+ \phi(x,t)|^2)\, dx = I_{11}(t) + I_{12}(t),$$

$$I_2(t) = \int\limits_{\Omega_{\delta - h\sqrt{d}/2}} (|\partial_t \nabla \phi(x,t)|^2 + |D \nabla_h^+ \phi(x,t)|^2)\, dx = I_{21}(t) + I_{22}(t)$$

and D is the differential operator defined as $(DF)_{ij} = \frac{\partial F_j}{\partial x_i}$, $1 \leq i, j \leq d$, for all vector functions $F = (F_1, \cdots, F_d)$.

In order to estimate the terms $I_{ij}(t)$, $1 \leq i, j \leq 2$, we apply a stationary phase argument (see [25], [50]) based on iterative integrations by parts and motivated by the fact that both $\partial_t \phi$ and $\nabla_h^+ \phi(x,t)$ given by (3.34) and (3.35) are oscillatory integrals. Observe that the phase $\psi(\eta\gamma h, x - x^*, t)$ does not depend only on the integration variable η as in the classical Stationary Phase Lemma setting, but also on the space and time variables x and t and on the product γh.

The oscillatory exponential $\exp(i\psi(\gamma h \eta, x - x^*, t)/h)$ satisfies the following identity:

$$\exp\Big(\frac{i}{h}\psi(\eta\gamma h, x - x^*, t)\Big) = \frac{1}{i\gamma} \frac{\nabla_\eta \psi(\eta\gamma h, x - x^*, t)}{|\nabla_\eta \psi(\eta\gamma h, x - x^*, t)|^2} \cdot \nabla_\eta \exp\Big(\frac{i}{h}\psi(\eta\gamma h, x - x^*, t)\Big). \tag{3.38}$$

Let us define the operator $\mathcal{L}$ as follows

$$\mathcal{L}\widehat{\sigma}(\eta, x - x^*, t) = \mathrm{div}_\eta\Big(\frac{\nabla_\eta \psi(\eta\gamma h, x - x^*, t)}{|\nabla_\eta \psi(\eta\gamma h, x - x^*, t)|^2} \widehat{\sigma}(\eta)\Big). \tag{3.39}$$

In what follows, we explain how the stationary phase argument works on $\partial_t \phi(x,t)$. By iterated integrations by parts and taking into account that $\widehat{\sigma} \in C_c^\infty(B^d(0,1))$ and then the boundary terms vanish, we obtain that for all $N \in \mathbb{N}^*$ the function $\partial_t \phi(x,t)$ can be transformed in the following way:

$$\partial_t \phi(x,t) = \frac{\gamma^{d/2}}{(2\pi)^d} \int\limits_{B^d(0,1)} \Big(-\frac{1}{i\gamma}\Big)^N \mathcal{L}^N \widehat{\sigma}(\eta, x - x^*, t) \exp\Big(\frac{i}{h}\psi(\gamma h \eta, x, t)\Big)\, d\eta. \tag{3.40}$$

Set v_d to be the volume of the d-dimensional unit ball. By the Cauchy-Schwartz inequality applied in (3.40), we obtain that for all $N \in \mathbb{N}^*$ the following inequality is valid:

$$|\partial_t \phi(x,t)|^2 \leq \frac{\gamma^{d-2N} v_d}{(2\pi)^{2d}} \int\limits_{B^d(0,1)} |\mathcal{L}^N \widehat{\sigma}(\eta, x - x^*, t)|^2\, d\eta. \tag{3.41}$$

The following result shows that for $N \in \mathbb{N}^*$ large enough, $\mathcal{L}^N \widehat{\sigma}(\eta, x - x^*, t)$ has good integrability properties not only in $\eta \in B^d(0,1)$, but also in $x \in \Omega_{\delta - h\sqrt{d}/2}$.

Lemma 3.3.3. *For all $N \in \mathbb{N}^*$, there exists a function $f_N((D^\alpha \widehat{\sigma}(\eta))_{|\alpha|\leq N}) \in L^1_\eta(B^d(0,1))$, depending also on T, δ, $(\|D^\alpha \omega_{d,1}\|_{L^\infty(B^d(\eta_0,\gamma h))})_{|\alpha|\leq N}$, but independent on x, such that*

$$|\mathcal{L}^N(\eta, x-x^*,t)|^2 \leq \frac{f_N((D^\alpha\widehat{\sigma}(\eta))_{|\alpha|\leq N})}{|x-x_h(t)|^{2N}}. \tag{3.42}$$

We postpone the proof of this lemma. From the inequalities (3.41) and (3.42), we see that for $N > d/2$ we have

$$\begin{aligned}
I_{11}(t) &= \int\limits_{\Omega_{\delta-h\sqrt{d}/2}(t)} |\partial_t\phi(x,t)|^2\,dx \leq \frac{\gamma^{d-2N}v_d}{(2\pi)^{2d}} \int\limits_{\Omega_{\delta-h\sqrt{d}/2}(t)}\int\limits_{B^d(0,1)} |\mathcal{L}^N(\eta,x-x^*,t)|^2\,d\eta\,dx \\
&\leq \frac{\gamma^{d-2N}v_d}{(2\pi)^{2d}} \int\limits_{\Omega_{\delta-h\sqrt{d}/2}(t)} \frac{1}{|x-x_h(t)|^{2N}}\,dx \int\limits_{B^d(0,1)} f_N((D^\alpha\widehat{\sigma}(\eta))_{|\alpha|\leq N})\,d\eta \\
&= \frac{\gamma^{d-2N}v_d}{(2\pi)^{2d}}\frac{dv_d}{2N-d}\frac{1}{(\delta-h\sqrt{d}/2)^{2N-d}} \int\limits_{B^d(0,1)} f_N((D^\alpha\widehat{\sigma}(\eta))_{|\alpha|\leq N})\,d\eta \\
&\leq C_N^{11}(T,\widehat{\sigma},\delta,\eta_0)\gamma^{-(2N-d)}.
\end{aligned} \tag{3.43}$$

In order to estimate $I^{12}(t)$, the same stationary phase argument should be applied d times with $\widehat{\sigma}(\eta)$ replaced with $\widehat{\sigma}(\eta)\widehat{g}^k_{12}(\eta)$, $1\leq k\leq d$, where

$$\widehat{g}^k_{12}(\eta) = \frac{\omega_1(\eta_{0,k}+\gamma h\eta_k)}{\omega_{d,1}(\eta_0+\eta\gamma h)}\exp\Big(\frac{i}{2}(\eta_{0,k}+\gamma h\eta_k)\Big).$$

Since $\omega_{d,1}\in C^\infty(\Pi^d_1\setminus\{0\})$ and $\eta_0\neq 0$, then for γh sufficiently small, $\widehat{g}^k_{12}\in C^\infty(B^d(0,1))$ and then $\widehat{\sigma}\widehat{g}^k_{12}\in C^\infty_c(B^d(0,1))$. Then we can conclude that there exists a constant $C^{12}_N(T,\widehat{\sigma},\delta,\eta_0)$ such that

$$I_{12}(t)\leq C^{12}_N(T,\widehat{\sigma},\delta,\eta_0)\gamma^{-(2N-d)}. \tag{3.44}$$

In order to estimate $I_{21}(t)$, observed firstly that $\partial_t\nabla\phi(x,t)$ has the following integral form:

$$\partial_t\nabla\phi(x,t) = \frac{\gamma^{d/2}}{(2\pi)^d}\int\limits_{B^d(0,1)} \widehat{\sigma}(\eta)\frac{i}{h}(\gamma h\eta+\eta_0)\exp\Big(\frac{i}{h}\psi(\gamma h\eta,x-x^*,t)\Big)\,d\eta. \tag{3.45}$$

We should apply the same stationary phase argument as we did with $I_{11}(t)$ d times, with $\widehat{\sigma}(\eta)$ replaced by $\widehat{\sigma}(\eta)\widehat{g}^k_{21}(\eta)$, $1\leq k\leq d$, with

$$\widehat{g}^k_{21}(\eta) = \frac{i}{h}(\gamma h\eta_k+\eta_{0,k}).$$

Observe that $\widehat{g}^k_{21}\in C^\infty(B^d(0,1))$, therefore $\widehat{\sigma}\widehat{g}^k_{21}\in C^\infty_c(B^d(0,1))$, for all $1\leq k\leq d$ and there exists a constant $C^{21}_N(T,\widehat{\sigma},\delta,\eta_0)$ such that

$$I_{21}(t)\leq \frac{1}{h^2}C^{21}_N(T,\widehat{\sigma},\delta,\eta_0)\gamma^{-(2N-d)}. \tag{3.46}$$

In order to estimate $I_{22}(t)$, observe firstly that the matrix $D\nabla^+_h\phi(x,t)$ has the following integral form:

$$D\nabla^+_h\phi(x,t) = \frac{\gamma^{d/2}}{(2\pi)^d}\int\limits_{B^d(0,1)} \widehat{\sigma}(\eta)\Big(\widehat{g}^i_{12}(\eta)\widehat{g}^j_{21}(\eta)\Big)_{1\leq i,j\leq d}\exp\Big(\frac{i}{h}\psi(\gamma h\eta,x-x^*,t)\Big)\,d\eta. \tag{3.47}$$

Due to the arguments before, we see that $\widehat{\sigma}\widehat{g}^i_{12}\widehat{g}^j_{21} \in C^\infty_c(B^d(0,1))$ for all $1 \le i,j \le d$. We apply d^2 times the stationary phase argument we applied for $I_{11}(t)$ with $\widehat{\sigma}(\eta)$ replaced by $\widehat{\sigma}(\eta)\widehat{g}^i_{12}(\eta)\widehat{g}^j_{21}(\eta)$, $1 \le i,j \le d$ and we conclude that there exists a constant $C^{22}_N(T,\widehat{\sigma},\delta,\eta_0)$ such that

$$I_{22}(t) \le \frac{1}{h^2} C^{22}_N(T,\widehat{\sigma},\delta,\eta_0)\gamma^{-(2N-d)}. \tag{3.48}$$

Using (3.43), (3.44), (3.46) and (3.48), the right hand side in (3.37) can be bounded from above as follows:

$$I_1(t)+2c(d)h(I_1(t)I_2(t))^{1/2} \le \big[C^1_N(T,\widehat{\sigma},\delta,\eta_0)+2c(d)\big(C^1_N(T,\widehat{\sigma},\delta,\eta_0)C^2_N(T,\widehat{\sigma},\delta,\eta_0)\big)^{1/2}\big]\gamma^{-(2N-d)},$$

with $C^i_N(T,\widehat{\sigma},\delta,\eta_0) = C^{i1}_N(T,\widehat{\sigma},\delta,\eta_0) + C^{i2}_N(T,\widehat{\sigma},\delta,\eta_0)$, $i = 1,2$. This concludes the proof of (3.30). □

Proof of Lemma 3.3.3. We argue by induction upon $N \in \mathbb{N}^*$.

Firstly, we check the inequality (3.42) for $N = 1$. Since $\mathrm{div}(Fg) = g\mathrm{div}F + F\cdot\nabla g$, for all vectorial function F and scalar one g, we deduce that

$$\mathcal{L}\widehat{\sigma}(\eta,x-x^*,t) = \nabla_\eta\widehat{\sigma}(\eta)\cdot\frac{\nabla_\eta\psi(\gamma h\eta,x-x^*,t)}{|\nabla_\eta\psi(\gamma h\eta,x-x^*,t)|^2} + \widehat{\sigma}(\eta)\mathrm{div}_\eta\Big(\frac{\nabla_\eta\psi(\gamma h\eta,x-x^*,t)}{|\nabla_\eta\psi(\gamma h\eta,x-x^*,t)|^2}\Big).$$

It is easy to verify that

$$\mathrm{div}_\eta\Big(\frac{\nabla_\eta\psi(\gamma h\eta,x-x^*,t)}{|\nabla_\eta\psi(\gamma h\eta,x-x^*,t)|^2}\Big) = \frac{\gamma h\Delta_\eta\psi(\gamma h\eta,x-x^*,t)}{|\nabla_\eta\psi(\gamma h\eta,x-x^*,t)|^2} - \\ - \frac{2\gamma h(\nabla_\eta\psi(\gamma h\eta,x-x^*,t),D^2_\eta\psi(\gamma h\eta,x-x^*,t)\nabla_\eta\psi(\gamma h\eta,x-x^*,t))}{|\nabla_\eta\psi(\gamma h\eta,x-x^*,t)|^4}.$$

We conclude that

$$|\mathcal{L}\widehat{\sigma}(\eta)| \le \frac{1}{|\nabla_\eta\psi(\gamma h\eta,x-x^*,t)|}\Big[|\widehat{\sigma}(\eta)|+\gamma h|\nabla_\eta\widehat{\sigma}(\eta)|\frac{|\Delta_\eta\psi(\gamma h\eta,x-x^*,t)|+2|D^2_\eta\psi(\gamma h\eta,x-x^*,t)|}{|\nabla_\eta\psi(\gamma h\eta,x-x^*,t)|}\Big].$$

Observe that $\nabla_\eta\psi(\eta,x,t) = t\nabla_\eta\omega_{d,1}(\eta_0+\eta)+x$. Therefore,

$$\nabla_\eta\psi(\gamma h\eta,x-x^*,t) = x - x_h(t) + t\gamma h\eta D^2\omega_{d,1}(\eta_0+\gamma h\eta'),\ \text{with}\ \eta' \in B^d(0,1) \tag{3.49}$$

Then $|\nabla_\eta\psi(\gamma h\eta,x-x^*,t)| \ge |x-x_h(t)| - T\gamma h\|D^2_\eta\omega_{d,1}\|_{L^\infty(B^d(\eta_0,\gamma h))}$. Since $\gamma h << 1$ and T is finite, for h small enough, we can guarantee that for all $x \in \Omega_{\delta-h\sqrt{d}/2}(t)$

$$T\gamma h\|D^2_\eta\omega_{d,1}\|_{L^\infty(B^d(\eta_0,\gamma h))} \le \frac{\delta-h\sqrt{d}/2}{2} \le \frac{1}{2}|x-x_h(t)|.$$

and that

$$|\nabla_\eta\psi(\gamma h\eta,x-x^*,t)| \ge \frac{1}{2}|x-x_h(t)|. \tag{3.50}$$

On the other hand, for all $\alpha \in \mathbb{N}^d$, with $|\alpha| = 2$, $D^\alpha_\eta\psi(\eta,x,t) = tD^\alpha\omega_{d,1}(\eta_0+\eta)$ does not depend on x. Then for all $x \in \Omega_{\delta-h\sqrt{d}/2}(t)$, we have that

$$\Big(|\widehat{\sigma}(\eta)| + \gamma h|\nabla_\eta\widehat{\sigma}(\eta)|\frac{|\Delta_\eta\psi(\gamma h\eta,x-x^*,t)|+2|D^2_\eta\psi(\gamma h\eta,x-x^*,t)|}{|\nabla_\eta\psi(\gamma h\eta,x-x^*,t)|}\Big)^2 \\ \le \Big(|\widehat{\sigma}(\eta)| + \gamma h|\nabla_\eta\widehat{\sigma}(\eta)|\frac{T\|\Delta_\eta\omega_{d,1}\|_{L^\infty(B^d(\eta_0,\gamma h))}+2T\|D^2_\eta\omega_{d,1}\|_{L^\infty(B^d(\eta_0,\gamma h))}}{\delta-h\sqrt{d}/2}\Big)^2 \\ =: \frac{1}{4}f_1((D^\alpha\widehat{\sigma}(\eta))_{|\alpha|\le 1}).$$

For the inductive step, since the multi-dimensional case is too technical, we restrict to the $1-d$ case, for which

$$\mathcal{L}\widehat{\sigma}(\eta, x-x^*, t) = \partial_\eta\Big(\widehat{\sigma}(\eta)\widehat{\theta}(\eta)\Big),$$

with

$$\widehat{\theta}(\eta) := \widehat{\theta}(\eta, x-x^*, t) = \frac{1}{\partial_\eta\psi(\eta_0+\gamma h\eta, x-x^*, t)}.$$

Set $\mathcal{A}_{N,j} := \{\alpha = (\alpha_0, \alpha_1, \cdots, \alpha_N) \in \mathbb{N}^{N+1} : \alpha_0+\alpha_1+\alpha_2\cdots+\alpha_N = N, \alpha_1+2\alpha_2+\cdots+N\alpha_N = N-j\}$

We assume that for each $\alpha \in \mathcal{A}_{N,j}$, there exists $a_\alpha \in \mathbb{R}$ depending only on α (and then only on N and j) such that $\mathcal{L}^N\widehat{\sigma}$ can be written in the following way:

$$\mathcal{L}^N\widehat{\sigma}(\eta, x-x^*, t) = \sum_{j=0}^{N}\widehat{\sigma}^{(j)}(\eta)\sum_{\alpha\in\mathcal{A}_{N,j}} a_\alpha(\widehat{\theta}(\eta))^{\alpha_0}(\widehat{\theta}^{(1)}(\eta))^{\alpha_1}\cdots(\widehat{\theta}^{(N)}(\eta))^{\alpha_N}. \tag{3.51}$$

It is clear that for $N=1$, the explicit formula (3.51) is valid, since $\mathcal{L}\widehat{\sigma}(\eta, x-x^*, t) = \widehat{\sigma}(\eta)\widehat{\theta}^{(1)}(\eta) + \widehat{\sigma}^{(1)}(\eta)\widehat{\theta}(\eta)$. The two possibilities for α are $\alpha = (0,1) \in \mathcal{A}_{1,0}$ and $\alpha = (1,0) \in \mathcal{A}_{1,1}$. Suppose that the explicit form (3.51) is valid for N and we prove it for $N+1$. Since

$$\mathcal{L}^{N+1}\widehat{\sigma}(\eta, x-x^*, t) = \widehat{\theta}^{(1)}(\eta)\mathcal{L}^N\widehat{\sigma}(\eta, x-x^*, t) + \widehat{\theta}(\eta)\partial_\eta\mathcal{L}^N\widehat{\sigma}(\eta, x-x^*, t),$$

we conclude that

$$\begin{aligned}
\mathcal{L}^{N+1}\widehat{\sigma}(\eta, x-x^*, t) &= \sum_{j=0}^{N}\widehat{\sigma}^{(j)}(\eta)\sum_{\alpha\in\mathcal{A}_{N,j}} a_\alpha(\widehat{\theta}(\eta))^{\alpha_0}(\widehat{\theta}^{(1)}(\eta))^{\alpha_1+1}(\widehat{\theta}^{(2)}(\eta))^{\alpha_2}\cdot\ldots\cdot(\widehat{\theta}^{(N)}(\eta))^{\alpha_N} \\
&+ \sum_{j=1}^{N+1}\widehat{\sigma}^{(j)}(\eta)\sum_{\alpha\in\mathcal{A}_{N,j-1}} a_\alpha(\widehat{\theta}(\eta))^{\alpha_0+1}(\widehat{\theta}^{(1)}(\eta))^{\alpha_1}\cdot\ldots\cdot(\widehat{\theta}^{(N)}(\eta))^{\alpha_N} \\
&+ \sum_{j=0}^{N}\widehat{\sigma}^{(j)}(\eta)\sum_{\alpha\in\mathcal{A}_{N,j}} a_\alpha\big[\alpha_0(\widehat{\theta}(\eta))^{\alpha_0}(\widehat{\theta}^{(1)}(\eta))^{\alpha_1+1}(\widehat{\theta}^{(2)}(\eta))^{\alpha_2}\cdot\ldots\cdot(\widehat{\theta}^{(N)}(\eta))^{\alpha_N} \\
&+ \alpha_1(\widehat{\theta}(\eta))^{\alpha_0+1}(\widehat{\theta}^{(1)}(\eta))^{\alpha_1-1}(\widehat{\theta}^{(2)}(\eta))^{\alpha_2+1}(\widehat{\theta}^{(3)}(\eta))^{\alpha_3}\cdot\ldots\cdot(\widehat{\theta}^{(N)}(\eta))^{\alpha_N} + \ldots \\
&+ \alpha_N(\widehat{\theta}(\eta))^{\alpha_0+1}(\widehat{\theta}^{(1)}(\eta))^{\alpha_1}\cdot\ldots\cdot(\widehat{\theta}^{(N-1)}(\eta))^{\alpha_{N-1}}(\widehat{\theta}^{(N)}(\eta))^{\alpha_N-1}\widehat{\theta}^{(N+1)}(\eta)\big].
\end{aligned}$$

Observe that if the explicit expression (3.51) would be true, the elements $\beta \in \mathcal{A}_{N+1,j}$ can be obtained from elements $\alpha \in \mathcal{A}_{N,i}$ by one of the following four procedures:

p1. For all $\alpha \in \mathcal{A}_{N,j}$, $0 \le j \le N$, one obtains a vector $\beta = (\alpha_0, \alpha_1+1, \alpha_2, \cdots, \alpha_N, 0) \in \mathcal{A}_{N+1,j}$.

p2. For all $\alpha \in \mathcal{A}_{N,j-1}$, $1 \le j \le N+1$, one obtains a vector $\beta = (\alpha_0+1, \alpha_1, \cdots, \alpha_N, 0) \in \mathcal{A}_{N+1,j}$.

p3. For all $\alpha \in \mathcal{A}_{N,j}$, $0 \le j \le N$ with $\alpha_k \neq 0$ and $1 \le k \le N-1$, one obtains a vector $\beta = (\alpha_0+1, \alpha_1, \cdots, \alpha_{k-1}, \alpha_k-1, \alpha_{k+1}+1, \alpha_{k+2}, \cdots, \alpha_N, 0) \in \mathcal{A}_{N+1,j}$.

p4. For all $\alpha \in \mathcal{A}_{N,j}$ with $\alpha_N \neq 0$, one obtains a vector $\beta = (\alpha_0+1, \alpha_1, \cdots, \alpha_{N-1}, \alpha_N-1, 1)$.

The iterative step consists in proving that for all $\beta \in \mathcal{A}_{N+1,j}$, $0 \le j \le N+1$, there exists an $\alpha \in \mathcal{A}_{N,i}$, for some $0 \le i \le N$, such that β can be obtained from α by one of the four procedures described above. Fix $\beta \in \mathcal{A}_{N+1,j}$, for some $0 \le j \le N+1$. Only one of the following situations is possible:

- If $\beta_1 \neq 0$ (which, from the two equations verified by the elements of $\mathcal{A}_{N+1,j}$, implies that $\beta_{N+1} = 0$ and $j \neq N+1$), we construct $\alpha = (\beta_0, \beta_1-1, \beta_2, \beta_N) \in \mathcal{A}_{N,j}$, so β is obtained from α by the procedure ($p1$).

- If $\beta_0 \neq 0$ and $1 \leq j \leq N+1$ (which implies that $\beta_{N+1} = 0$), we construct $\alpha = (\beta_0 - 1, \beta_1, \cdots, \beta_N) \in \mathcal{A}_{N,j-1}$, such that β is obtained from α by the procedure $(p2)$.
- If for some $2 \leq k \leq N$ $\beta_k \neq 0$ (which implies that $\beta_0 \neq 0$ and $\beta_{N+1} = 0$), we construct $\alpha = (\beta_0 - 1, \beta_1, \cdots, \beta_{k-1} + 1, \beta_k - 1, \beta_{k+1}, \cdots, \beta_N) \in \mathcal{A}_{N,j}$, such that β is obtained from α by the procedure $(p3)$.
- If $\beta_{N+1} = 1$ (which implies that $j = 0$, $\beta_0 = N$ and $\beta_k = 0$ for all $1 \leq k \leq N$), we construct $\alpha = (N, 0 \cdots, 0, 1) \in \mathcal{A}_{N,0}$, such that β is obtained from α by the procedure $(p4)$.

We conclude that the inductive step is verified and that $\mathcal{L}^N \widehat{\sigma}$ has the precise formula (3.51). Observe that $\widehat{\theta}(\eta)$ can be written as a composition of two functions: $\widehat{\theta} = \widehat{f} \circ \widehat{g}(\eta)$, where $\widehat{f}(\eta) = 1/\eta$ and $\widehat{g}(\eta) = \partial_\eta \psi(\gamma h \eta, x - x^*, t)$. For $k \geq 1$, we will use the following formula to compute the derivative of order k of $\widehat{\theta}$ (cf. [3]):

$$\frac{1}{k!}(\widehat{f} \circ \widehat{g})^{(k)}(\eta) = \sum_{1 \leq i \leq k} \frac{1}{i!} \widehat{f}^{(i)}(\widehat{g}(\eta)) \sum_{k_1 + \cdots + k_i = k} \frac{\widehat{g}^{(k_1)}(\eta)}{k_1!} \cdot \ldots \cdot \frac{\widehat{g}^{(k_i)}(\eta)}{k_i!}.$$

We take into account that, in our particular case,

$$\widehat{f}^{(i)}(\eta) = (-1)^i i! \eta^{-(i+1)} \text{ and } \widehat{g}^{(i)}(\eta) = t(\gamma h)^i \omega_1^{(i+1)}(\eta_0 + \gamma h \eta).$$

Therefore, we obtain that

$$\widehat{\theta}^{(k)}(\eta) = (\gamma h)^k \sum_{1 \leq i \leq k} \frac{k!(-t)^i}{(\partial_\eta \psi(\gamma h \eta, x - x^*, t))^{i+1}} \sum_{k_1 + \cdots + k_i = k} \frac{\omega_1^{(k_1+1)}(\eta_0 + \gamma h \eta)}{k_1!} \cdot \ldots \cdot \frac{\omega_1^{(k_i+1)}(\eta_0 + \gamma h \eta)}{k_i!}.$$

Using the inequality (3.50) and the fact that $x \in \Omega_{\delta - h/2}(t)$ in the last identity, we obtain that

$$|\theta(\eta)| \leq \frac{1}{|x - x_h(t)|} \text{ and } |\theta^{(k)}(\eta)| \leq \frac{(\gamma h)^k}{|x - x_h(t)|} c_k(T, \delta, \eta_0), \tag{3.52}$$

with

$$c_k(T, \delta, \eta_0) = 2k! \sum_{1 \leq i \leq k} \Big(\frac{2T}{\delta - h/2} \Big)^i \sum_{k_1 + \cdots + k_i = k} \frac{\|\omega_1^{(k_1+1)}\|_{L^\infty(B(\eta_0, \gamma h))}}{k_1!} \cdot \ldots \cdot \frac{\|\omega_1^{(k_i+1)}\|_{L^\infty(B(\eta_0, \gamma h))}}{k_i!}.$$

Using (3.52) in (3.51), we conclude the proof of (3.42), with

$$f_N((D^\alpha \widehat{\sigma}(\eta))_{|\alpha| \leq N}) = \Big(\sum_{j=0}^{N} |\widehat{\sigma}^{(j)}(\eta)| (\gamma h)^{N-j} \sum_{\alpha \in \mathcal{A}_{N,j}} |a_\alpha| (c_1(T, \delta, \eta_0))^{\alpha_1} \cdot \ldots \cdot (c_N(T, \delta, \eta_0))^{\alpha_N} \Big)^2.$$

□

Set $\mathcal{S}(\mathbb{R}^d)$ to be the class of the Schwartz functions on $\mathbb{R}^d$. The following result extends Theorem 3.3.1 to the class of Schwartz functions.

Theorem 3.3.2. *Fix $T > 0$ and consider a wave number $\eta_0 = h\xi_0 \in \Pi_1^d \setminus \{0\}$ and a point $x^* \in B^d(0,1)$ such that the semi-discrete GCC (3.27) is not verified. Consider $\widehat{\sigma} \in \mathcal{S}(\mathbb{R}^d)$ and $\gamma = \gamma(h) > 0$, $r = r(h) > 0$ to be two functions such that the following conditions are verified*

$$\lim_{h \to 0} \gamma(h) = +\infty, \quad \lim_{h \to 0} \frac{r(h)}{h\gamma(h)} = \infty \text{ and } \lim_{h \to 0} r(h) = 0. \tag{3.53}$$

In the semi-discrete wave equation (3.3) the initial data $(\overrightarrow{\phi}^{h,0}, \overrightarrow{\phi}^{h,1})$ *are such that their SDFTs verify the following two requirements:*

$$\widehat{\phi}^{h,0}(\xi) = \gamma^{-d/2}\widehat{\sigma}(\gamma^{-1}(\xi-\xi_0))\frac{\exp(-i\xi\cdot x^*)}{i\omega_{d,h}(\xi)}\chi_{B^d(\xi_0,r/h)} \text{ and } \widehat{\phi}^{h,1}(\xi) = i\omega_{d,h}(\xi)\widehat{\phi}^{h,0}(\xi). \quad (3.54)$$

Then for all $\alpha > 0$*, there exists a constant* $c_\alpha = c_\alpha(T, \widehat{\sigma}, \eta_0) > 0$ *not depending on* h *such that the discrete observability constant* $C_h(T)$ *in (3.12) satisfies* $C_h(T) \geq c_\alpha\gamma^\alpha$.

Proof. For simplicity, set $\widetilde{r} := r/(\gamma h)$. One of the characteristics of the Schwartz functions is the fact that $D^\alpha\widehat{\sigma} \in L^p(\mathbb{R}^d)$ for all $1 \leq p \leq \infty$. One of the consequences of this fact and of the fact that $\widetilde{r} >> 1$ is that the total energy of the solutions of (3.3) with initial data (3.54) tends to a constant:

$$E_h(\overrightarrow{\phi}^{h,0}, \overrightarrow{\phi}^{h,1}) = \frac{1}{(2\pi)^d}\int\limits_{B^d(0,\widetilde{r})} |\widehat{\sigma}(\eta)|^2\, d\eta \sim \frac{1}{(2\pi)^d}\int\limits_{\mathbb{R}^d} |\widehat{\sigma}(\eta)|^2\, d\eta.$$

The proof is also based on the stationary phase like arguments and we will mention here only the differences with respect to the estimates for the term $I_1(t)$ defined by (3.37). Instead of a time derivative of the numerical solution given by (3.34), now the time derivative $\partial_t\phi(x,t)$ is given by

$$\partial_t\phi(x,t) = \frac{\gamma^{d/2}}{(2\pi)^d}\int\limits_{B^d(0,\widetilde{r})} \widehat{\sigma}(\eta)\exp\Big(\frac{i}{h}\psi(\gamma h\eta, x-x^*, t)\Big)\, d\eta, \quad (3.55)$$

with the phase ψ defined by (3.32).

Consider $\widehat{\rho} \in C_c^\infty(B^d(0,1))$, such that $\widehat{\rho} \equiv 1$ in $B^d(0,1/2)$ and $0 \leq \widehat{\rho} \leq 1$ in $B^d(0,1)$. We define the following two functions:

$$A_i(x,t) = \frac{\gamma^{d/2}}{(2\pi)^d}\int\limits_{B^d(0,\widetilde{r})} \widehat{\sigma}(\eta)\widehat{\rho}\Big(\frac{\eta}{\widetilde{r}}\Big)\exp\Big(\frac{i}{h}\psi(\gamma h\eta, x-x^*, t)\Big)\, d\eta \text{ and } A_o(x,t) = \partial_t\phi(x,t) - A_i(x,t).$$

Since $\widehat{\sigma}\widehat{\rho}(\cdot/\widetilde{r}) \in C_c^\infty(B^d(0,\widetilde{r}))$, we can apply the stationary phase argument on $A_i(x,t)$.

The condition $r << 1$**.** We have to guarantee that (3.50) is verified. We have seen that due to the identity (3.49), this is guaranteed by the fact that $|t\gamma h\eta D^2\omega_{d,1}(\eta_0 + \gamma h\eta')|$, with $\eta' \in B^d(0,\widetilde{r})$, can be done arbitrarily small for all $\eta \in B^d(0,\widetilde{r})$. In our case, this is true since $\gamma h\eta, \gamma h\eta' \in B^d(0,r)$, $r << 1$ and $t \leq T$, with T being finite. Then

$$|t\gamma h\eta D^2\omega_{d,1}(\eta_0 + \gamma h\eta')| \leq Tr\|D^2\omega_{d,1}\|_{L^\infty(B^d(\eta_0,r))}.$$

The condition $\widetilde{r} >> 1$**.** Within the stationary phase argument applied on $A_i(x,t)$, we have to compute the derivatives of any order of $\widehat{\sigma}\widehat{\rho}(\cdot/\widetilde{r})$. Using the Leibniz rule, we can compute explicitly the derivative os order $\alpha \in \mathbb{N}^d$:

$$D^\alpha\Big(\widehat{\sigma}\widehat{\rho}\Big(\frac{\cdot}{\widetilde{r}}\Big)\Big)(\eta) = \sum_{\beta=0}^{\alpha}\binom{\alpha}{\beta}D^{\alpha-\beta}\widehat{\sigma}(\eta)\widetilde{r}^{-|\beta|}D^\beta\widehat{\rho}\Big(\frac{\eta}{\widetilde{r}}\Big).$$

Since $\widetilde{r} >> 1$, we have that $\widetilde{r}^{-|\beta|} << 1$ for all $\beta \neq 0$. So, by multiplying the profile $\widehat{\sigma}$ by $\widehat{\rho}(\cdot/\widetilde{r})$, we do not loose powers of h when apply the stationary phase argument.

Another motivation to consider $\widetilde{r} >> 1$ arises when we evaluate the error term $A_o(x,t)$. Indeed, by the Parseval identity and the hypothesis on $\widehat{\rho}$, we get:

$$\begin{aligned}\int\limits_{\Omega_{\delta-h\sqrt{d}/2}(t)} |A_o(x,t)|^2\, dx \leq \int_{\mathbb{R}^d} |A_o(x,t)|^2\, dx &= \frac{1}{(2\pi)^d}\int\limits_{B^d(0,\widetilde{r})} |\widehat{\sigma}(\eta)|^2\Big|1-\widehat{\rho}\Big(\frac{\eta}{\widetilde{r}}\Big)\Big|^2\, d\eta \\ &\leq \frac{1}{(2\pi^d)}\int\limits_{\mathbb{R}^d\setminus B^d(0,\widetilde{r}/2)} |\widehat{\sigma}(\eta)|^2\, d\eta. \quad (3.56)\end{aligned}$$

Due to the fact that $\widehat{\sigma} \in \mathcal{S}(\mathbb{R}^d)$ and that $\widetilde{r} >> 1$, the right hand side in (3.56) converges to zero in a polynomial manner at arbitrary order in the variable $\widetilde{r}^{-1}$. □

This example of wave packet clearly illustrates the classical effect due to the group velocity that is plotted in Figure 3.3 below. We plot in red the initial velocity $\phi^1(x) = \exp(-\gamma|x|^2/2)\exp(i\xi_0 x)$ (we also take $i|\xi|\widehat{\phi^0} = \widehat{\phi^1}$), with $h = 0.005$, $\xi_0 h = 19\pi/20$, $\gamma = h^{-\alpha}$, $2\alpha = 0.75$ in $d = 1$ (left) and $d = 2$ (right), which coincides, up to an exponentially small error (with respect to h), with the discrete initial data (3.54). In blue, we plot the time derivative of the solution of the continuous wave equation (3.1) and in black the one corresponding to the semi-discrete wave equation (3.3) with initial data (3.54) at time $t = 1$. For $d = 2$, we represent only the projection of these (continuous and discrete) solutions on the x-plane. This experiment shows that, as theory predicts, the semi-discrete wave packets propagate with a velocity that is much smaller than the one of the continuous wave equation.

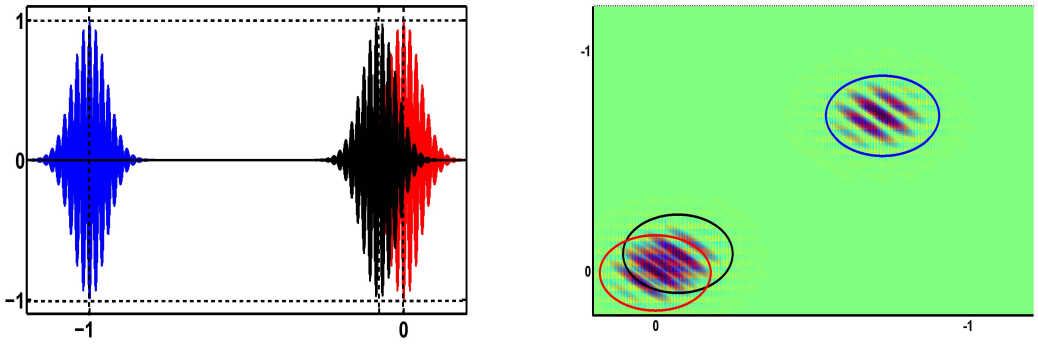

Figure 3.3: Transport of continuous versus discrete high frequency wave packets in dimension $d = 1$ and $d = 2$.

3.3.2 Fully discrete finite difference schemes for the wave equation.

In this subsection, we show that for the fully discrete finite difference schemes for the wave equation the associated discrete observability constant is expected to blow-up polynomially at any order by adapting the construction of highly oscillatory wave packet in the first section. For simplicity, we reduce to the $1-d$ case. Consider $k, h > 0$, $\lambda = k/h \leq 1$ and an uniform grid of $\mathbb{R} \times \mathbb{R}_+$ given by the points $(x_j, t_n)_{j\in\mathbb{Z}, n\in\mathbb{N}} = (jh, nk)_{j\in\mathbb{Z}, n\in\mathbb{N}}$. Under these assumptions, the following Cauchy problem associated to the fully discrete finite difference approximation of the $1-d$ wave equation is considered:

$$\begin{cases} \frac{\phi_j^{n+1} - 2\phi_j^n + \phi_j^{n-1}}{k^2} = \frac{\phi_{j+1}^n - 2\phi_j^n + \phi_{j-1}^n}{h^2}, & j \in \mathbb{Z}, n \in \mathbb{N}\setminus\{0,1\} \\ \phi_j^0 = \psi_j^0, \phi_j^1 = \psi_j^0 + k\psi_j^1, & j \in \mathbb{Z}. \end{cases} \tag{3.57}$$

By applying the SDFT at the scale k in the space variable in (3.57), we obtain the recurrence

$$\begin{cases} \widehat{\phi}^{h,n+1}(\xi) - 2\widehat{\phi}^{h,n}(\xi) + \widehat{\phi}^{h,n-1}(\xi) = -4\lambda^2 \sin^2\left(\frac{\xi h}{2}\right)\widehat{\phi}^{h,n}(\xi) & \xi \in \Pi_h, n \in \mathbb{N}\setminus\{0,1\} \\ \widehat{\phi}^{h,0}(\xi) = \widehat{\psi}^{h,0}(\xi), \quad \widehat{\phi}^{h,1}(\xi) = \widehat{\psi}^{h,0}(\xi) + k\widehat{\psi}^{h,1}(\xi), & \xi \in \Pi_h, \end{cases} \tag{3.58}$$

where $\widehat{\phi}^{h,n}(\xi)$ is the SDFT of the sequence $\overrightarrow{\phi}^n = (\phi_j^n)_{j\in\mathbb{Z}}$. Set

$$\omega_{h,k}(\xi) := \frac{2}{k}\arcsin\left(\lambda \sin\left(\frac{\xi h}{2}\right)\right).$$

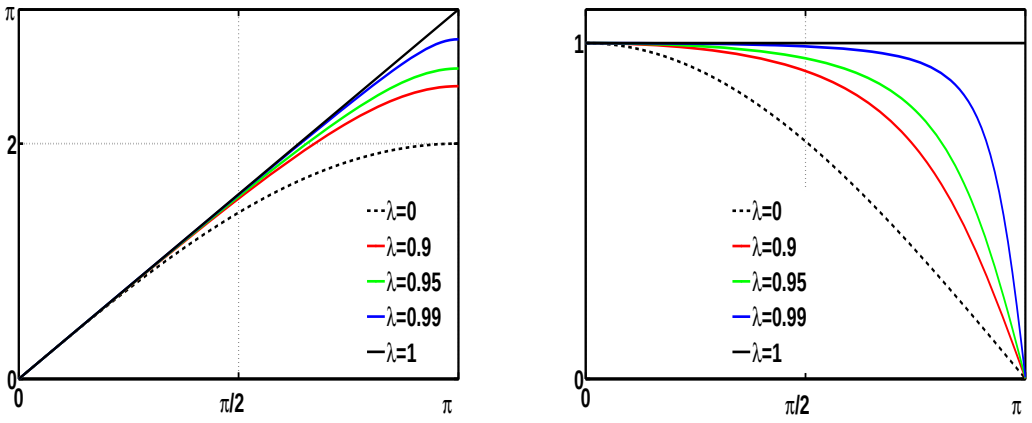

Figure 3.4: The dispersion relation $\omega_{1,\lambda}(\xi)$ for the fully discrete finite difference $1-d$ wave equation versus the corresponding group velocity.

The solution of (3.58) is given by

$$\widehat{\phi}^{h,n}(\xi) = \widehat{\psi}^{h,0}(\xi)\frac{\cos\left(t_{n-1/2}\omega_{h,k}(\xi)\right)}{\cos\left(t_{1/2}\omega_{h,k}(\xi)\right)} + k\widehat{\psi}^{h,1}(\xi)\frac{\sin\left(t_n\omega_{h,k}(\xi)\right)}{\sin\left(t_1\omega_{h,k}(\xi)\right)}.$$

The solution of (3.57) is given by means of the inverse SDFT:

$$\phi_j^n = \frac{1}{2\pi}\int\limits_{\Pi_h}\left[\widehat{\psi}^{h,0}(\xi)\frac{\cos\left(t_{n-1/2}\omega_{h,k}(\xi)\right)}{\cos\left(t_{1/2}\omega_{h,k}(\xi)\right)} + k\widehat{\psi}^{h,1}(\xi)\frac{\sin\left(t_n\omega_{h,k}(\xi)\right)}{\sin\left(t_1\omega_{h,k}(\xi)\right)}\right]\exp(i\xi x_j)\,d\xi.$$

The group velocity associated to $\omega_{h,k}(\xi)$ is given by

$$\partial_\xi\omega_{h,k}(\xi) = \frac{\cos\left(\frac{\xi h}{2}\right)}{\sqrt{1-\lambda^2\sin^2\left(\frac{\xi h}{2}\right)}}.$$

In [44], pp. 81, the total energy associated to (3.57) has been introduced in the following manner

$$E_{h,k}(\overrightarrow{\psi}^{h,0},\overrightarrow{\psi}^{h,1}) = \frac{h}{2}\sum_{j\in\mathbb{Z}}\left|\frac{\phi_j^{n+1}-\phi_j^n}{k}\right|^2 + \frac{h}{2}\sum_{j\in\mathbb{Z}}\frac{\phi_{j+1}^{n+1}-\phi_j^{n+1}}{h}\frac{\phi_{j+1}^n-\phi_j^n}{h}$$

and is shown that it is conserved in $n\in\mathbb{N}$. We are interested in the following discrete version of the observability inequality

$$E_{h,k}(\overrightarrow{\psi}^{h,0},\overrightarrow{\psi}^{h,1}) \leq C_{h,k}(T)\frac{hk}{2}\sum_{t_n\in[0,T]}\sum_{|x_j|>1}\left[\left|\frac{\phi_j^{n+1}-\phi_j^n}{k}\right|^2 + \sum_{|x_j|>1}\left|\frac{\phi_{j+1}^{n+1}-\phi_j^{n+1}}{h}\frac{\phi_{j+1}^n-\phi_j^n}{h}\right|\right] \tag{3.59}$$

More precisely, we are interested for the behavior of the In Figure 3.4, observe that for $\lambda=1$, the group velocity is 1, for all $\xi\in\Pi_h$, whereas for $\lambda\in(0,1)$ the group velocity vanishes for $\xi=\pi/h$. Therefore, for $\lambda\in(0,1)$ there exist discrete waves propagating arbitrarily slow and the construction in Section 1 can be adapted to prove that the constant $C_{h,k}(T)$ blows-up at least polynomially as $h\to 0$.

3.4 Qualitative properties of Gaussian wave packets

In this section, we focus on the $1-d$ semi-discrete version of the Cauchy problem (3.3). In the initial data (3.54), consider the particular case when

$$\widehat{\sigma}(\xi) = \widehat{\sigma}_\gamma(\xi) = \sqrt{\frac{2\pi}{\gamma}} \exp\Big(- \frac{\xi^2}{2\gamma^2}\Big). \tag{3.60}$$

In this section, we analyze some qualitative properties of the solution to the semi-discrete wave equation (3.3) corresponding to initial data (3.54) and (3.60). More precisely, we will focus on the dispersive behavior and symmetry of the solution with respect to the principal ray of Geometric Optics. For simplicity, consider $\delta = 1 - x^*$ and

$$\gamma = h^{-\alpha}, \alpha > 0.$$

Dispersion analysis. In this paragraph, we will analyze the values of α for which the solution $\overrightarrow{\phi}^h(t)$ of (3.3) with initial data (3.54) resembles to a Gaussian function in the physical space. When this holds, we analyze the amplitude and the essential support of this Gaussian profile. Since we deal with phenomena aimed to be observable from a computational point of view for example, we will work in the $\ell^\infty(h\mathbb{Z})$ framework. We get the following estimate

Proposition 3.4.1. *Set $\theta(x,t) := \exp(i\xi_0 x + it\omega_h(\xi_0))$, $\theta_j(t) := \theta(x_j,t)$. $y_j(t) := x_j - x^* + t\omega_1'(\eta_0)$,*

$$\gamma^2(t) := \frac{\gamma^2}{1 - ith\gamma^2\omega_1''(\eta_0)},$$

$$\phi_{approx}(x,t) := \theta(x,t)\sqrt{\frac{\gamma^2(t)}{\gamma^2}}\sigma_{\gamma(t)}(x - x(t))$$

and

$$\phi_{approx,j}(t) := \phi_{approx}(x_j,t) = \theta_j(t)\sqrt{\frac{\gamma^2(t)}{\gamma^2}}\sigma_{\gamma(t)}(y_j(t)).$$

For all $t \in \mathbb{R}_+$, the solution of (3.3) corresponding to the initial data (3.54) verifies the following estimate as $h \to 0$:

$$\|\overrightarrow{\phi}^h(t) - \overrightarrow{\phi}^h_{approx}(t)\|_{\ell^\infty(h\mathbb{Z})} = O\big(th^2\gamma^3\big). \tag{3.61}$$

Remark 3.4.1. *Before proceeding to the proof of Proposition 3.4.1, we do some observations concerning the amplitude of $\overrightarrow{\phi}^h_{approx}(t)$ and on the right hand side of (3.61).*

- *For an uniform time t as $h \to 0$, the right hand side in (3.61) is of order $h^{2-3\alpha}$. According to this estimate, we know for sure that $\overrightarrow{\phi}^h_{approx}(t)$ is a good approximation of $\overrightarrow{\phi}^h(t)$ at least for all $\alpha < 2/3$ and for any finite t.*

- *Observe that $|\phi_{approx,j}(t)| = \sqrt{|\frac{\gamma^2(t)}{\gamma^2}|}\sigma_{Re(\gamma^2(t))}(y_j(y))$. Moreover, for all t finite, but sufficiently large,*

$$Re(\gamma^2(t)) = \frac{1}{h^{4\alpha}\delta^2 + t^2h^2\delta^{-2}|\omega_1''(\eta_0)|^2} \sim \begin{cases} h^{-4\alpha}, & \alpha \le 1/2 \\ h^{-2}, & \alpha \in (1,2) \end{cases}$$

and

$$|\frac{\gamma^2(t)}{\gamma^2}| = \frac{1}{\sqrt{1 + t^2h^2\gamma^4|\omega_1''(\eta_0)|^2}} \to \begin{cases} 1, & \alpha < 1/2 \\ \frac{1}{\sqrt{1+t^2\delta^{-4}|\omega_1'(\eta_0)|^2}}, & \alpha = 1/2 \\ 0, & \alpha > 1/2 \end{cases} \quad \text{as } h \to 0.$$

We conclude that:

1. For $\alpha \in (0, 1/2)$ the support of $\overrightarrow{\phi}^h_{approx}(t)$ is of order $h^{2\alpha-1}$ for any finite t, but $\|\overrightarrow{\phi}^h_{approx}(t)\|_{\ell^\infty(h\mathbb{Z})} \to 1$ as $h \to 0$.

2. For $\alpha = 1/2$, the support is $h^{-1/4}$ and $\|\overrightarrow{\phi}^h_{approx}(t)\|_{\ell^\infty(h\mathbb{Z})} \to \frac{1}{\sqrt{1+t^2\delta^{-4}|\omega_1''(\eta_0)|^2}}$ as $h \to 0$.

3. For $\alpha \in (1/2, 1)$, the support of $\overrightarrow{\phi}^h_{approx}(t)$ changes from order $h^{\alpha/2-1}$ at $t = 0$ to $h^{-\alpha/2}$ in a finite time t, but $\|\overrightarrow{\phi}^h_{approx}(t)\|_{\ell^\infty(h\mathbb{Z})} \to 0$ as $h \to 0$.

Proof of Proposition 3.4.1. By decomposing $\omega_1(\eta)$ in its Taylor series as

$$\omega_1(\eta) = \omega_1(\eta_0) + (\eta - \eta_0)\omega_1'(\eta_0) + (\eta - \eta_0)^2\omega_1''(\eta_0)/2 + (\eta - \eta_0)^3\omega_1'''(\eta')/3!,$$

with η' lying on the segment joining η and η_0, we obtain that

$$\phi_j(t) - \theta_j(t)\sqrt{\frac{\gamma^2(t)}{\gamma}}\,\sigma_{\gamma(t)}(y_j(t)) = \theta_j(t)\big(-D_j^1(t) + D_j^2(t)\big),$$

with

$$D_j^1(t) = \frac{1}{2\pi}\int\limits_{\mathbb{R}\setminus\Pi_h} \widehat{\sigma}_\gamma(\xi - \xi_0)\exp(i(\xi - \xi_0)y_j(t))\exp\big(it(\xi - \xi_0)^2 h\omega_1''(\eta_0)/2\big)\,d\xi$$

and

$$D_j^2 = \frac{1}{2\pi}\int\limits_{\Pi_h} \widehat{\sigma}_\gamma(\xi - \xi_0)\exp(i(\xi - \xi_0)y_j(t))\big(\exp\big(it(\xi - \xi_0)^3h^2\omega_1'''(\xi' h)/3!\big) - 1\big)\,d\xi,$$

with $\eta' = \xi' h$.

By an easy computation, taking into account that $\xi_0 \in (0, \pi/h)$,

$$|D_j^1(t)| \le \frac{1}{2\pi}\int\limits_{\mathbb{R}\setminus\Pi_h} |\widehat{\sigma}_\gamma(\xi - \xi_0)|\,d\xi \le \sqrt{2}\exp\big(-\frac{(\pi - \eta_0)^2}{4\gamma^2h^2}\big),$$

so $\|\overrightarrow{D}^1\|_{\ell^\infty(h\mathbb{Z})}$ is exponentially small as $h \to 0$ if $\eta_0 \in (0, \pi)$. On the other hand,

$$|D_j^2(t)| \le \frac{1}{2\pi}\int\limits_{\mathbb{R}} \widehat{\sigma}_\gamma(\xi)\Big|2\sin\Big(\frac{t\xi^3h^2\omega_1'''(\xi' h)}{3!2}\Big)\Big|\,d\xi.$$

Taking into account the fact that $\|\omega_1'''\|_{L^\infty(\Pi_1)} = 1/4$ and that

$$\int\limits_{\mathbb{R}} \exp(-|\xi|^2/\gamma^2)|\xi|^k\,d\xi = \gamma^{k+1}\Gamma\Big(\frac{k+1}{2}\Big), \tag{3.62}$$

we obtain that

$$\|\overrightarrow{D}^2(t)\|_{\ell^\infty(h\mathbb{Z})} \le \frac{th^2\gamma^3}{3\sqrt{2\pi}},$$

which conclude the proof of the estimate (3.61). □

Lack of symmetry of the numerical concentrated waves. We also observe a lack of symmetry of the wave packet $\overrightarrow{\phi}^h(t)$ defined by (3.54) for $\alpha \ge 1/2$. In the following, we make more rigorous this property. In fact, we have to study the symmetry with respect to the discrete

ray $x(t)$. This means to evaluate the numerical solution at points $x(t) \pm x$, with $x \in \mathbb{R}$. This kind of evaluations requires to introduce the following function defined for all $x \in \mathbb{R}$:

$$\phi(x,t) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{\sigma}_\gamma(\xi - \xi_0) \exp(-i(\xi - \xi_0)x^*) \exp(it\omega_h(\xi)) \exp(i\xi x)\, d\xi. \tag{3.63}$$

The following estimate shows that at the grid points $\phi(\cdot,t)$ is exponentially close to $\overrightarrow{\phi}^h(t)$:

$$\sup_{j\in\mathbb{Z}} |\phi_j(t) - \phi(x_j,t)| \leq \frac{1}{2\pi} \int_{\mathbb{R}\setminus\Pi_h} |\widehat{\sigma}_\gamma(\xi - \xi_0)|\, d\xi \leq \sqrt{2\pi} \exp\Big(- \frac{(\pi - \eta_0)^2}{4\gamma^2 h^2}\Big).$$

It is not difficult to see that $\phi(x_j,t) - \phi_j(t)$ is also exponentially small in the energy norm. The previous estimate shows that the analysis of the lack of symmetry for $\phi(\cdot,t)$ is transferred modulo an exponential error to $\overrightarrow{\phi}^h(t)$. The same analysis in the proof of Proposition 3.4.1 shows that

$$\|\phi(\cdot,t) - \phi_{approx}(\cdot,t)\|_{L^\infty(\mathbb{R})} = O\big(th^2\gamma^3\big). \tag{3.64}$$

From the last estimate, we deduce that, for a finite t, $\alpha \leq 2/3$ and all $x \in \mathbb{R}$,

$$|\phi(x(t)+x,t) - \phi(x(t)-x,t)| \sim |\phi_{approx}(x(t)+x,t) - \phi_{approx}(x(t)-x,t)|.$$

Proposition 3.4.2. *For ϕ_{approx} introduced in Proposition 3.4.1, $\forall x \in \mathbb{R}$ and $\forall t \geq 0$, the following identity holds:*

$$|\phi_{approx}(x(t)+x,t) - \phi_{approx}(x(t)-x,t)| = 2|\sin(\xi_0 x)| \sqrt{\Big|\frac{\gamma(t)}{\gamma}\Big|}\sigma_{Re(\gamma(t))}(x). \tag{3.65}$$

Proof of Proposition 3.4.2. Observe that $\phi_{approx}(x,t)$ is the product of two functions: $\theta(x,t)$ which is not symmetric with respect to $x(t)$ and $\sqrt{\frac{\gamma^2(t)}{\gamma^2}}\sigma_{\gamma(t)}(x - x(t))$ which is symmetric with respect to $x_h(t)$. Then

$$|\phi_{approx}(x(t)+x,t) - \phi_{approx}(x(t)-x,t)| = |\theta(x(t)+x,t) - \theta(x(t)-x,t)|\Big|\sqrt{\frac{\gamma^2(t)}{\gamma^2}}\Big||\sigma_{\gamma(t)}(x)|$$

and (3.65) follows by taking into account the definition of θ and the identity $|\sigma_{\gamma(t)}(x)| = \sigma_{\mathrm{Re}(\gamma(t))}(x)$. □

3.5 Discrete WKB expansions

In the following, we analyze more precisely the behavior of these wave packets close to the selected ray of Geometric Optics. To do this, we approximate the analytic dispersion relation $\omega_{d,h}(\xi)$ by a polynomial one.

The time derivative of the solution of (3.3) with initial data (3.54) is given by the following wave packet

$$\partial_t \phi_{\mathbf{j}}(t) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \sqrt{\frac{2\pi}{\gamma}}^d \widehat{\phi}(\xi/\gamma)\chi_{B^d(0,r/h)}(\xi) \exp(it\omega_{d,h}(\xi + \xi_0)) \exp(i\xi\cdot(x_{\mathbf{j}} - x^*))\, d\xi. \tag{3.66}$$

To simplify the presentation and, without loss of generality, we set $\xi_0 = 0$ and $x^* = 0$. Since $\widehat{\phi} \in \mathcal{S}(\mathbb{R}^d)$, we may also neglect the characteristic function in (3.66). Setting $\omega(\xi) = \omega_{d,1}(\xi + \eta_0)$ and $x = x_{\mathbf{j}} - x^*$, the wave packet (3.66) can be written as

$$u(x,t) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \sqrt{\frac{2\pi}{\gamma}}^d \widehat{\phi}\big(\frac{\xi}{\gamma}\big) \exp\Big(\frac{it}{h}\omega(\xi h)\Big) \exp(i\xi\cdot x)\, d\xi. \tag{3.67}$$

In view of the analyticity of $\omega(\eta)$, we can split it as $\omega(\eta) = L(\eta) + D(\eta) + R(\eta)$, where $L(\eta) = \omega(0) + \nabla\omega(0)\cdot\eta$ is the linear part of ω, D is the second order term in the Taylor expansion of ω about $\eta = 0$ and R is the corresponding reminder, given explicitly by

$$D(\eta) = \sum_{|\alpha|=2} \frac{1}{\alpha!} D^\alpha\omega(0)\eta^\alpha, R(\eta) = \sum_{|\alpha|=3} \frac{3}{\alpha!}\eta^\alpha \int_0^1 (1-\lambda)^2 D^\alpha\omega(\lambda\eta)\, d\lambda.$$

Factoring it out the time dependent complex exponential generated by the zero order term in the Taylor expansion of the dispersion relation, we may decompose u as $u\exp(-it\omega_{d,1}(0)/h) = v + v^R$, where

$$v(x,t) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \sqrt{\frac{2\pi}{\gamma}}^d \widehat{\phi}(\frac{\xi}{\gamma}) \exp\Big(\frac{it}{h} D(\xi h)\Big) \exp(i\xi\cdot(x + t\nabla\omega(0)))\, d\xi$$

and

$$v^R(x,t) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \sqrt{\frac{2\pi}{\gamma}}^d \widehat{\phi}(\frac{\xi}{\gamma}) \exp\Big(\frac{it}{h} D(\xi h)\Big)\Big(\exp\Big(\frac{it}{h} R(\xi h)\Big) - 1\Big) \exp(i\xi\cdot(x + t\nabla\omega(0)))\, d\xi.$$

Our main result determines the conditions on γ and $\widehat{\phi}$, such that v^R is a reminder term.

Theorem 3.5.1. *For all $t > 0$ and $\widehat{\phi} \in \mathcal{S}(\mathbb{R}^d)$, the following estimate holds:*

$$\frac{\|v^R(\cdot,t)\|^2_{L^2(\mathbb{R}^d)}}{\|u(\cdot,t)\|^2_{L^2(\mathbb{R}^d)}} \leq C(\widehat{\phi}) h^4 \gamma^6 t^2, \tag{3.68}$$

where $C(\widehat{\phi}) = C\||\cdot|^3\widehat{\phi}\|^2_{L^2(\mathbb{R}^d)}/\|\widehat{\phi}\|^2_{L^2(\mathbb{R}^d)}$ and $C = \Big(\sum_{|\alpha|=3} \frac{1}{\alpha!}\|D^\alpha\omega\|_{L^\infty(B(0,h\gamma))}\Big)^2$

For a finite time interval $t \in [0,T]$, one can guarantee that v^R is small in the sense of (3.68) when

$$\gamma h^{2/3} << 1, \tag{3.69}$$

which is a more restrictive condition on γ than $\gamma h << 1$ in (3.28), guaranteing that the energy concentrated outside the ray is polynomially small. Indeed, (3.68) yields an asymptotic description of the solution globally in space-time, i.e. $u \sim \exp(it\omega(0)/h)v$. Intuitively, the scale (3.69) is motivated by the fact that $|R(\xi h)/h| \leq \sqrt{C}h^2|\xi|^3$. For this to be asymptotically small, the width of the Fourier transform of the profile has to be limited by (3.69).

The function v is a solution of the following PDE (cf. [40]):

$$\partial_t v = \nabla\omega(0)\cdot\nabla_x v - ih\sum_{|\alpha|=2} D^\alpha\omega(0) D_x^\alpha v. \tag{3.70}$$

This is a transport equation perturbed by an asymptotically small (as h tends to zero) Schrödinger like second order term. The following result emphasizes that there exists solutions of (3.70) for which the relevant scale is

$$\gamma h^{1/2} = 1.$$

Theorem 3.5.2. *There exist solutions of (3.70) admitting the following asymptotic expansion:*

$$v(x,t) = \sum_{j=0}^{\infty} h^{j/2} a_j\left(\frac{x}{h^{1/2}}, \frac{t}{h^{1/2}}\right) \exp\Big(ix\cdot\frac{\eta_0}{h^{1/2}} + it\omega_2\Big(\frac{\eta_0}{h^{1/2}}\Big)\Big), \tag{3.71}$$

where $\eta_0 \in \mathbb{R}^d$ is a fixed wave number, $\omega_2(\eta) = \nabla\omega(0)\cdot\eta + h\sum_{|\alpha|=2}\frac{1}{\alpha!}D^\alpha\omega(0)\eta^\alpha$ is the dispersion relation corresponding to (3.70) and $(a_j(x,t))_{j\in\mathbb{N}}$ solve the following system of PDEs:

$$\partial_t a_0 = \nabla\omega(0)\cdot\nabla_x a_0 \text{ and } \partial_t a_{j+1} = \nabla\omega(0)\cdot\nabla_x a_{j+1} - i\sum_{|\alpha|=2}\frac{1}{\alpha!}D^\alpha\omega(0)\sum_{0\le\beta<\alpha}(i\eta_0)^\beta D_x^{\alpha-\beta}a_j, \forall j\in\mathbb{N}.$$

By taking the support of the initial datum of a_0 to be compact, one can observe that v is concentrated along a neighborhood of the ray of width $\sqrt{h}$. This scale $\sqrt{h}$ is critical due to the added dispersion that the Schrödinger like term introduces. Once the asymptotic expansion of v is given as in (3.71), one immediately gets that of u and therefore of $\partial_t\overrightarrow{\phi}$ and of $\overrightarrow{\phi}$. In this way, we get the asymptotic form of the high frequency wave packet.

Proof. The proof is based on the so-called Wentzel-Kramer-Brillouin (WKB) asymptotic method (cf. [48]). For all $\epsilon > 0$ and $\eta_0 \in \mathbb{R}^d$, fixed, the plane wave $\exp(ix\cdot\eta_0/\epsilon + it\omega_2(\eta_0/\epsilon))$ is a solution of (3.70). In order to concentrate in space, we multiply this plane wave by a profile $a(x,t)$. In this way, we consider solutions of (3.70) of the form

$$v(x,t) = a(x,t)\exp(ix\cdot\eta_0/\epsilon + it\omega_2(\eta_0/\epsilon)), \text{with } a(x,t) = \sum_{j=0}^{\infty}\epsilon^j\widetilde{a}_j(x,t).$$

The profile $a(x,t)$ satisfies the equation

$$\partial_t a = \nabla\omega(0)\cdot\nabla_x a - ih\sum_{|\alpha|=2}\frac{1}{\alpha!}D^\alpha\omega(0)\sum_{0\le\beta<\alpha}\left(\frac{i\eta_0}{\epsilon}\right)^\beta D_x^{\alpha-\beta}a.$$

For any $j\in\mathbb{N}$, as well as x and t are at the same scale, $\partial_t\widetilde{a}_j$ and $\nabla\omega(0)\cdot\nabla_x\widetilde{a}_j$ are of the same order. Nevertheless, in the term involving the second-order derivatives of $\widetilde{a}_j$,

$$-ih\sum_{|\alpha|=2}\frac{1}{\alpha!}D^\alpha\omega(0)\sum_{0\le\beta<\alpha}\left(\frac{i\eta_0}{\epsilon}\right)^\beta D_x^{\alpha-\beta}\widetilde{a}_j,$$

two scales are involved: h and h/ϵ. When ϵ is small, $h << h/\epsilon$. The simplest case to find the functions $\widetilde{a}_j$ is when they satisfy a linear system of PDEs, such that the knowledge of $\widetilde{a}_j$ determines directly the knowledge of $\widetilde{a}_{j+1}$. This is possible when $\widetilde{a}_j(x,t) = a_j(\gamma x,\gamma t)$. The functions a_j are related as follows:

$$\sum_{j=0}^{\infty}\Big[\gamma\epsilon^j\big(\partial_t a_j - \nabla\omega(0)\cdot\nabla_x a_j\big) + ih\sum_{|\alpha|=2}\frac{1}{\alpha!}D^\alpha\omega(0)\sum_{0\le\beta<\alpha}(i\eta_0)^\beta\frac{\gamma^{|\alpha-\beta|}}{\epsilon^{|\beta|}}D_x^{\alpha-\beta}a_j\Big] = 0.$$

When $|\beta|$ increases by one, $|\alpha-\beta|$ decreases by one. The choice $\gamma = \epsilon^{-1}$ is the only one for which

$$\frac{\gamma^{k_1}}{\epsilon^{k_2}} = \frac{\gamma^{k_1-1}}{\epsilon^{k_2+1}}, \forall k_1,k_2\in\mathbb{N}.$$

With this choice, for all $0\le\beta<\alpha$, with $|\alpha|=2$, all second order terms are at the same scale. Indeed,

$$\frac{\gamma^{|\alpha-\beta|}}{\epsilon^{|\beta|}} = \frac{1}{\epsilon^{|\alpha-\beta|+|\beta|}} = \frac{1}{\epsilon^{|\alpha|}} = \frac{1}{\epsilon^2}.$$

We find the relation between ϵ and h by imposing that the scale of $\partial_t a_{j+1} - \nabla\omega(0)\cdot\nabla_x a_{j+1}$ to be the same as the one accompanying $\sum_{|\alpha|=2}\frac{1}{\alpha!}D^\alpha\omega(0)\sum_{0\le\beta<\alpha}(i\eta_0)^\beta D_x^{\alpha-\beta}a_j$, for all $j\in\mathbb{N}$. This yields, $\gamma\epsilon^{j+1} = \epsilon^j h\epsilon^{-2}$, i.e. $\epsilon = h^{1/2}$ and $\gamma = h^{-1/2}$. □

This kind of expansion can be further developed, incorporating higher order terms of the Taylor expansion of the dispersion relation and a multiple-scale ansatz. This issue has to be understood better and will be developed in detail in the future.

3.6 Semi-discrete Gaussian beams

In this subsection, we will adapt the construction of the Gaussian beams developed for the case of hyperbolic operators in [46] and then adapted for the continuous wave equation with singular coefficients in [38]. The difficulty to adapt such constructions to the case of numerical approximations is that the discrete operators depend implicitly on a small parameter h and then one has to limit the range of the parameter ϵ in the Gaussian beam ansatz according to h. We will see that the correct scale is $\epsilon \sim h$.

Since the bi-characteristic rays associated to the semi-discretization (3.3) are continuous in the space variable, we will extend the discrete Laplacian $\triangle_h$ to all the real Euclidian space $\mathbb{R}^d$ in the following way

$$\triangle_h f(x) = \frac{1}{h^2}\sum_{k=1}^{d}\big(f(x+h\mathbf{e}_k) - 2f(x) + f(x-h\mathbf{e}_k)\big). \tag{3.72}$$

For fixed $h>0$, consider the wave equation associated to the extended operator $\triangle_h$:

$$\Box_h \phi(x,t) = \partial_t^2\phi(x,t) - \triangle_h\phi(x,t) = 0, \quad x\in\mathbb{R}^d, t>0. \tag{3.73}$$

The bi-characteristic rays associated to the extended wave operator $\Box_h$ are the same as for the operator extended to the grid points: $(x(s),t(s),\xi(s),\tau(s)) = (x^* - 2\omega_{d,h}(\xi_0)\nabla\omega_{d,h}(\xi_0)s, 2\tau_0 s, \xi_0, \tau_0)$, with $|\tau_0|^2 = \omega_{d,h}^2(\xi_0)$ and $\omega_{d,h}$ defined by (3.18). The only difference is that for the continuous operator $\Box_h$ the wave numbers ξ_0 are in $\mathbb{R}^d$, whereas for the discrete operator $\Box_h$, the wave numbers are from Π_h^d.

Consider the following Geometric Optics ansatz for the solutions of (3.73):

$$\phi(x,t) = a(x,t)\exp(i\psi(x,t)/\epsilon), \quad a(x,t) = \sum_{j=1}^{N} a_j(x,t)\epsilon^j. \tag{3.74}$$

The second-order time derivative of ϕ is

$$\partial_t^2\phi(x,t) = \exp(i\psi(x,t)/\epsilon)\Big[\partial_t^2 a(x,t) + \frac{2i}{\epsilon}\partial_t a(x,t)\partial_t\psi(x,t) + \frac{i}{\epsilon}a(x,t)\partial_t^2\psi(x,t) - \frac{1}{\epsilon^2}a(x,t)\big(\partial_t\psi(x,t)\big)^2\Big].$$

In order to organize in a suitable way the second-order discrete derivatives $\partial_{h,k}^2$, we will use the following identities

$$\partial_{h,k}^2(fg)_{\mathbf{j}} = f_{\mathbf{j}}\partial_{h,k}^2 g_{\mathbf{j}} + \partial_{h,k}^+ f_{\mathbf{j}}\partial_{h,k}^+ g_{\mathbf{j}} + \partial_{h,k}^- \mathbf{f}_j\partial_{h,k}^- g_{\mathbf{j}} + g_{\mathbf{j}}\partial_{h,k}^2 f_{\mathbf{j}},$$

$$\partial_{h,k}^+ + \partial_{h,k}^- = 2\partial_{h,k}, \partial_{h,k}^+ - \partial_{h,k}^- = h\partial_{h,k}^2$$

and, for all $a,b\in\mathbb{C}$,

$$\exp(ia) + \exp(-ib) = 2\cos\big(\frac{a+b}{2}\big)\exp\big(\frac{i(a-b)}{2}\big),$$

$$\exp(ia) - \exp(-ib) = 2i\sin\big(\frac{a+b}{2}\big)\exp\big(\frac{i(a-b)}{2}\big)$$

and

$$2 - \exp(ia) - \exp(-ib) = 4\sin^2\big(\frac{a+b}{2}\big) - 4i\cos\big(\frac{a+b}{2}\big)\sin\big(\frac{a-b}{4}\big)\exp\big(\frac{i(a-b)}{4}\big).$$

Then

$$\begin{aligned}\partial_{h,k}^2\phi(x,t) = \exp(i\psi(x,t)/\epsilon)\Big[&\frac{i}{\epsilon}\Big(2\partial_{h,k}a(x,t)\frac{\epsilon}{h}\sin\big(\partial_{h,k}\psi(x,t)\frac{h}{\epsilon}\big)\exp\big(\frac{ih^2}{2\epsilon}\partial_{h,k}^2\psi(x,t)\big)\\ &+ a(x,t)\cos\big(\partial_{h,k}\psi(x,t)\frac{h}{\epsilon}\big)\frac{4\epsilon}{h^2}\sin\big(\frac{h^2}{4\epsilon}\partial_{h,k}^2\psi(x,t)\big)\exp\big(\frac{ih^2}{4\epsilon}\partial_{h,k}^2\psi(x,t)\big)\Big)\\ &+ \partial_{h,k}^2 a(x,t)\cos\big(\partial_{h,k}\psi(x,t)\frac{h}{\epsilon}\big)\exp\big(\frac{ih^2}{2\epsilon}\partial_{h,k}^2\psi(x,t)\big) - \frac{1}{\epsilon^2}\frac{4\epsilon^2}{h^2}\sin^2\big(\frac{h}{2\epsilon}\partial_{h,k}\psi(x,t)\big)\Big].\end{aligned}$$

Summarizing, we get

$$\Box_h \phi(x,t) = \exp(i\psi(x,t)/\epsilon)\Big[\mathcal{R}^0 a(x,t) + \frac{i}{\epsilon}\mathcal{R}^1 a(x,t) - \frac{1}{\epsilon^2} a(x,t) r(x,t)\Big], \tag{3.75}$$

with $\mathcal{R}^0$ and $\mathcal{R}^1$ operators defined by

$$\mathcal{R}^0 f(x,t) = \partial_t^2 f(x,t) - \sum_{k=1}^{d} \cos\big(\partial_{h,k}\psi(x,t)\frac{h}{\epsilon}\big) \exp\big(\frac{ih^2}{2\epsilon}\partial_{h,k}^2\psi(x,t)\big)\partial_{h,k}^2 f(x,t),$$

$$\mathcal{R}^1 f(x,t) = 2\partial_t f(x,t)\partial_t\psi(x,t) - 2\frac{\epsilon}{h}\sum_{k=1}^{d} \partial_{h,k} f(x,t) \sin\big(\partial_{h,k}\psi(x,t)\frac{h}{\epsilon}\big) \exp\big(\frac{ih^2}{2\epsilon}\partial_{h,k}^2\psi(x,t)\big)$$

$$+ f(x,t)\Big(\partial_t^2\psi(x,t) - \frac{4\epsilon}{h^2}\sum_{k=1}^{d} \cos\big(\partial_{h,k}\psi(x,t)\frac{h}{\epsilon}\big) \sin\big(\frac{h^2}{4\epsilon}\partial_{h,k}^2\psi(x,t)\big) \exp\big(\frac{ih^2}{4\epsilon}\partial_{h,k}^2\psi(x,t)\big)\Big)$$

and

$$r(x,t) = (\partial_t\psi(x,t))^2 - \frac{4\epsilon^2}{h^2}\sum_{k=1}^{d} \sin^2\big(\frac{h}{2\epsilon}\partial_{h,k}\psi(x,t)\big).$$

There are three possibilities for h/ϵ:

- If $h/\epsilon \to \infty$, $r(x,t) \to (\partial_t\psi(x,t))^2$ and the term containing ϵ^{-2} cannot vanish on the bi-characteristic ray.
- If $h/\epsilon \to 0$, $r(x,t) \to (\partial_t\psi(x,t))^2 - \sum_{k=1}^{d}(\partial_{x_k}\psi(x,t))^2$. This is the principal symbol of the continuous wave equation evaluated in $(\nabla_x\psi(x,t), \partial_t\psi(x,t))$. The construction of Gaussian beams is based on solving the corresponding Eikonal equation or the principal symbol up to an arbitrary order on the ray. Then the case $h/\epsilon \to 0$ is not good since we do not obtain the semi-discrete Eikonal operator accompanying ϵ^{-2}.
- If $h/\epsilon \to 1$, $r(x,t) \to (\partial_t\psi(x,t))^2 - \sum_{k=1}^{d} 4\sin^2(\partial_{x_k}\psi(x,t)/2)$, which is the semi-discrete Eikonal operator.

For simplicity, consider $\epsilon = h$.

One of the main ideas in the Gaussian beam construction is to design Taylor series for the phase ψ and the amplitude a along the curve $(x(s), t(s))$. For this, one has to choose the partial derivatives of ψ and a on the bi-characteristic ray. Observe that the Eikonal r contains also the discrete partial derivatives of ψ w.r.t. the space variable x. We replace then the Eikonal $r(x,t)$ by

$$\widetilde{r}(x,t) = (\partial_t\psi(x,t))^2 - 4\sum_{k=1}^{d} \sin^2\big(\frac{1}{2}\partial_{x_k}\psi(x,t)\big). \tag{3.76}$$

We can do this since the error $r - \widetilde{r}$ is small

$$r(x,t) - \widetilde{r}(x,t) = 4\sum_{k=1}^{d} \sin\Big(\frac{1}{2}\big(\partial_{h,k}\psi(x,t) - \partial_{x_k}\psi(x,t)\big)\Big) \sin\Big(\frac{1}{2}\big(\partial_{h,k}\psi(x,t) + \partial_{x_k}\psi(x,t)\big)\Big) \tag{3.77}$$

$$\sim h^2\frac{1}{3}\sum_{k=1}^{d} \sin(\partial_{x_k}\psi(x,t))\partial_{x_k}^3\psi(x,t).$$

First step: Find the phase $\psi(x,t)$. This can be done by solving the Eikonal equation $p(x,t,\nabla_x\psi(x,t),\psi_t(x,t)) = 0$ up to an arbitrary order on the ray, where $p(x,t,\xi,\tau) = \tau^2 - 4\sum_{k=1}^{d}\sin^2(\xi_k/2)$. Observe that p is the principal symbol corresponding to (3.73) normalized to $h=1$. The rays of Geometric Optics corresponding to this symbol solve (3.19). The coefficients in problem (3.73) are constants, so the symbol p is independent of x and t. To simplify the notation, we will denote the symbol by $p = p(\xi,\tau)$.

Lemma 3.6.1 (Solving the Eikonal equation $\widetilde{r}(x,t)=0$ up to an arbitrary order on the ray)**.** *For all $R\in\mathbb{N}$, there exists a phase $\psi(x,t)=\psi_R(x,t)$ s.t. for all $m\in\mathbb{N}^{d+1}$, $|m|\le R$*

$$\frac{\partial^{|m|}\widetilde{r}(x(s),t(s))}{\partial t^{m_0}\partial x_1^{m_1}\cdots\partial x_d^{m_d}} = 0,$$

$$Im(\psi(x(s),t(s))) = 0 \text{ and } \Big(Im\Big(\frac{\partial^2\psi}{\partial x_i\partial x_j}(x(s),t(s))\Big)\Big)_{1\le i,j\le d} \text{ is positive definite.}$$

Proof of Lemma 3.6.1. The proof is the same as the one done for a similar result in [46]. It is based on the fact that

$$\widetilde{r}(x,t) = p(\nabla_x\psi(x,t),\partial_t\psi(x,t)).$$

By choosing $\nabla_x\psi(x(s),t(s)) = \xi_0$, $\partial_t\psi(x(s),t(s))$ and $M(s) = D^2_{x,t}\psi(x(s),t(s))$ solution of the Riccati equation $\dot{M}(s) + M(s)CM(s) = 0$, with $C = \mathrm{diag}(-2\cos(\xi_{0,1}),\cdots,-2\cos(\xi_{0,d}),2)$, one can find an initial data $M(0)$ s.t. $M(s)$ is symmetric and $(M_{i,j}(s))_{1\le i,j\le d}$ has the imaginary part positive definite for any $s\ge 0$. For $|\alpha|\ge 3$, $D^\alpha_{x,t}\psi(x(s),t(s))$ can be found by solving systems of first order linear ODEs. □

Second step: Find the amplitude $a(x,t)$. Remark in (3.74) that the amplitude a is a sum involving $N+1$ functions $a_j(x,t)$, $0\le j\le N$. First of all, we determine the coefficients for the different powers of h in $\Box_h\phi(x,t)$ according to the functions a_j. Set

$$r_j^0(x,t) := \mathcal{R}^0 a_j(x,t), \quad r_j^1(x,t) := \mathcal{R}^1 a_j(x,t),$$

with $\mathcal{R}^0$, $\mathcal{R}^1$ the operators introduced in (3.75). Taking into account (3.75) and (3.77), we obtain

$$\begin{aligned}\Box_h\phi(x,t) = \exp(i\psi(x,t)/h)\Big[\sum_{j=0}^{N} h^j\Big(r_j^0(x,t) - a_j(x,t)\frac{r(x,t)-\widetilde{r}(x,t)}{h^2}\Big)\\ + i\sum_{j=-1}^{N-1} h^j r_{j+1}^1(x,t) - \sum_{j=-2}^{N-2} h^j a_{j+2}(x,t)\widetilde{r}(x,t)\Big].\end{aligned}$$

Since the Eikonal equation $\widetilde{r}(x,t)=0$ was solved up to order R on the ray, the last term involving expressions of the form $a_j(x,t)\widetilde{r}(x,t)$ in the above identity will not contribute in the computation of $a_j(x,t)$. r_j^1 will definitely contribute in the determination of a_j.

As done for the phase ψ, we are interested in determining the Taylor series of a_j on the ray up to some order. Observe that r_j^1 contains only discrete derivatives in space and it would be more difficult to solve some system concerning the continuous derivatives of r_j^1 when discrete derivatives are involved. We first introduce a perturbation of $r_j^1(x,t)$ involving only continuous derivatives of a_j and of the phase ψ:

$$\begin{aligned}\widetilde{r}_j^1(x,t) = 2\partial_t a_j(x,t)\partial_t\psi(x,t) - 2\sum_{k=1}^{d}\partial_{x_k}a_j(x,t)\sin(\partial_{x_k}\psi(x,t))\\ + a_j(x,t)\Big(\partial_t^2\psi(x,t) - \sum_{k=1}^{d}\cos(\partial_{x_k}\psi(x,t))\partial_{x_k}^2\psi(x,t)\Big).\end{aligned} \tag{3.78}$$

Indeed, by the well known error estimate

$$|\partial_{h,k}f(x)-\partial_{x_k}f(x)|\le\frac{h^2}{3!}\|\partial^3_{x_k}f\|_\infty,$$

for all $f\in C^3_b(\mathbb{R}^d)$, we have

$$r^1_j(x,t)-\widetilde{r}^1_j(x,t)=h\widehat{r}^1_j(x,t)+h^2\breve{r}^1_j(x,t), \tag{3.79}$$

where

$$\begin{aligned}\widehat{r}^1_j(x,t)=&-2\sum_{k=1}^d\partial_{x_k}a_j(x,t)\sin(\partial_{x_k}\psi(x,t))\frac{1}{h}\Big(\exp\big(\frac{ih}{2}\partial^2_{h,k}\psi(x,t)\big)-1\Big)\\&-a_j(x,t)\sum_{k=1}^d\cos(\partial_{x_k}\psi(x,t))\partial^2_{x_k}\psi(x,t)\frac{1}{h}\Big(\exp\big(\frac{ih}{4}\partial^2_{h,k}\psi(x,t)\big)-1\Big)\end{aligned}$$

and

$$\begin{aligned}\breve{r}^1_j(x,t)=&-2\sum_{k=1}^d\frac{1}{h^2}\Big(\partial_{h,k}a_j(x,t)\sin(\partial_{h,k}\psi(x,t))-\partial_{x_k}a_j(x,t)\sin(\partial_{x_k}\psi(x,t))\Big)\exp\big(\frac{ih}{2}\partial^2_{h,k}\psi(x,t)\big)\\&-a_j(x,t)\sum_{k=1}^d\frac{1}{h^2}\Big(\cos(\partial_{h,k}\psi(x,t))\frac{4}{h}\sin\Big(\frac{h}{4}\partial^2_{h,k}\psi(x,t)\Big)-\cos(\partial_{x_k}\psi(x,t))\partial^2_{x_k}\psi(x,t)\Big)\exp\big(\frac{ih}{4}\partial^2_{h,k}\psi(x,t)\big).\end{aligned}$$

Then

$$\Box_h\phi(x,t)=\exp(i\psi(x,t)/h)\big[\sum_{j=-1}^{N+1}h^jr_j(x,t)-\sum_{j=-2}^{N-2}h^ja_{j+2}(x,t)\widetilde{r}(x,t)\big], \tag{3.80}$$

where

$$r_j(x,t)=r^0_j(x,t)-a_j(x,t)\frac{r(x,t)-\widetilde{r}(x,t)}{h^2}+i\widetilde{r}^1_{j+1}(x,t)+i\widehat{r}^1_j(x,t)+i\breve{r}^1_{j-1}(x,t),$$

for $-1\le j\le N+1$, $\breve{r}^1_{-2}(x,t)=\widehat{r}^1_{-1}(x,t)=\breve{r}^1_{-1}(x,t)=r^0_{-1}(x,t)=a_{-1}(x,t)=0$ and $r^0_{N+1}(x,t)=a_{N+1}(x,t)=\widetilde{r}^1_{N+1}(x,t)=\widetilde{r}_{N+2}(x,t)=\widehat{r}^1_{N+1}(x,t)=0$.

To find $a_j(x,t)$, $j=0,N$, we will solve $r_j(x,t)=0$ on the ray $(x(s),t(s))$ up to some order β_j, for $-1\le j\le N-1$, with β_j to be determined later on and depending on R.

Lemma 3.6.2 (Estimates on the discrete energy of the Gaussian Beam)**.** *For all $t\ge0$, there exists a constant $C(t)$ which does not depend on h s.t.*

$$E_h(\overrightarrow{\phi}^h(t),\partial_t\overrightarrow{\phi}^h(t))\sim C(t)h^{d/2-2},$$

where $\phi_j(t)=\phi(x_j,t)$ and $\phi(x,t)$ is the function given by the ansatz (3.74).

Proof of Lemma 3.6.2. The quantities involved in the discrete energy are

$$\begin{aligned}\partial^+_{h,k}\phi(x_{\mathbf{j}},t)=\frac{1}{2h}\exp(i\psi(x_{\mathbf{j}},t)/h)\Big[&h\partial^+_{h,k}a(x_{\mathbf{j}},t)\big(1+\exp(i\partial^+_{h,k}\psi(x_{\mathbf{j}},t))\big)\\&+\big(a(x_{\mathbf{j}}+h\mathbf{e}_k,t)+a(x_{\mathbf{j}},t)\big)\big(\exp(i\partial^+_{h,k}\psi(x_{\mathbf{j}},t))-1\big)\Big]\end{aligned}$$

and

$$\partial_t\phi(x_{\mathbf{j}},t)=\frac{1}{h}\exp(i\psi(x_{\mathbf{j}},t)/h)\big[h\partial_ta(x_{\mathbf{j}},t)+ia(x_{\mathbf{j}},t)\partial_t\psi(x_{\mathbf{j}},t)\big].$$

Since the time component of the bi-characteristic rays, $t(s) = 2\tau_0 s$, is linear, for any $t \in \mathbb{R}_+$, there exists an unique $s \in \mathbb{R}_+$ s.t. $t = t(s)$. Observe that the phase ψ can be written as

$$\psi(x,t(s)) = \psi(x(s),t(s)) + (x-x(s))\cdot\nabla_x\psi(x(s),t(s)) + \frac{1}{2}(x-x(s))D_x^2\psi(x(s),t(s))(x-x(s)) + \widetilde{\psi}(x,t(s)),$$

with

$$\widetilde{\psi}(x,t(s)) = \sum_{|\alpha|\geq 3}\frac{1}{\alpha!}(x-x(s))^\alpha D_x^\alpha\psi(x(s),t(s)).$$

Set $N(s) := \mathrm{Im}\big(D_x^2\psi(x(s),t(s))\big)$ and

$$y(x,s)\sqrt{h} := x - x(s). \tag{3.81}$$

We have seen that $\psi(x(s),t(s)) \in \mathbb{R}$ and $\nabla_x\psi(x(s),t(s)) = \xi(s) = \xi_0 \in \mathbb{R}^d$. Then

$$|\exp(i\psi(x_{\mathbf{j}},t(s))/h)| = \exp(-y(x_{\mathbf{j}},s)N(s)y(x_{\mathbf{j}},s))|\exp(i\widetilde{\psi}(x(s)+\sqrt{h}y(x_{\mathbf{j}},s),t(s))/h)|,$$

with

$$\frac{\widetilde{\psi}(x(s)+\sqrt{h}y(x_{\mathbf{j}},s),t(s))}{h} = \sum_{|\alpha|\geq 3}\frac{1}{\alpha!}(y(x_{\mathbf{j}},s))^\alpha h^{|\alpha|/2-1}D_x^\alpha\psi(x(s),t(s)).$$

For any $\alpha \in \mathbb{R}^d$ s.t. $|\alpha| \geq 3$ we have $|\alpha|/2 - 1 \geq 1/2$.

Using the above properties of $\psi(x,t)$ and we obtain that as $h \to 0$

$$\frac{E_h(\overrightarrow{\phi}(t),\partial_t\overrightarrow{\phi}(t))}{h^{d/2}-2} \to \frac{\pi^{d/2}}{(\det(N(s)))^{1/2}}|a_0(x(s),t(s))|^2\omega_{d,1}^2(\xi_0) \neq 0.$$

Then there exists two constants $C^\pm(s) > 0$ s.t., for all $t \geq 0$ and s s.t. $t = t(s)$, we have

$$C^-(s)h^{d/2-2} \leq E_h(\overrightarrow{\phi}(t),\partial_t\overrightarrow{\phi}(t)) \leq C^+(s)h^{d/2-2}.$$

□

Lemma 3.6.3. *Consider R, the order at which the Eikonal $\widetilde{r}$ vanishes on the ray $(x(s),t(s))$, and N, the number of terms in the Gaussian beam ansatz (3.74), s.t. $R - 2N = 2$. Then, for $\beta_k = R - 4 - 2k$, $-1 \leq k \leq N-1$, there exists a constant $C(N,t)$ s.t.*

$$\frac{\|\Box_h\overrightarrow{\phi}^h(t)\|^2_{\ell^2(h\mathbb{Z}^d)}}{E_h(\overrightarrow{\phi}^h(t),\partial_t\overrightarrow{\phi}^h(t))} \leq C(N,t)h^{R-1}.$$

Proof of Lemma 3.6.3. Firstly observe that using (3.80)

$$\|\Box_h\overrightarrow{\phi}^h(t)\|^2_{\ell^2(h\mathbb{Z}^d)} \leq (N+3)\sum_{k=-1}^{N+1}I_k^1(t) + (N+1)\sum_{k=-2}^{N-2}I_k^2(t),$$

with

$$I_k^1(t) = h^{d+2k}\sum_{\mathbf{j}\in\mathbb{Z}^d}|\exp(i\psi(x_{\mathbf{j}},t)/h)|^2|r_k(x_{\mathbf{j}},t)|^2$$

and

$$I_k^2(t) = h^{d+2k}\sum_{\mathbf{j}\in\mathbb{Z}^d}|\exp(i\psi(x_{\mathbf{j}},t)/h)|^2|a_{k+1}(x_{\mathbf{j}},t)|^2|\widetilde{r}(x_{\mathbf{j}},t)|^2.$$

The fact that the Eikonal $\widetilde{r}(x,t)$ vanishes on the ray up to order R can be written as

$$\widetilde{r}(x,t(s)) = \sum_{|\alpha|=R+1} \frac{1}{\alpha!}(x-x(s))^\alpha D_x^\alpha \widetilde{r}(x',t(s)),$$

where x' lies on the segment $(x,x(s))$. By the change of variable (3.81), we have

$$\frac{I_k^2(t)}{h^{d/2+2k+R+1}} \to |a_{k+1}(x(s),t(s))|^2 \int_{\mathbb{R}^d} \exp(-yN(s)y)\Big|\sum_{|\alpha|=R+1}\frac{1}{\alpha!}y^\alpha D_x^\alpha \widetilde{r}(x(s),t(s))\Big|^2 dy.$$

Then there exists a constant $c_k^2(s)>0$ s.t.

$$I_k^2(t) \le c_k^2(s) h^{d/2+2k+R+1}, \forall -2 \le k \le N-2.$$

Using the same arguments, if $r_k(x,t)$ vanishes up to order β_k, for $-1 \le k \le N-1$, we find that there exists a constant $c_k^1(s)>0$ s.t.

$$I_k^1(t) \le c_k^1(s) h^{d/2+2k+\beta_k+1}, \forall -1 \le k \le N-1.$$

For $k = N, N+1$, since $r_k(x,t)$ does not vanishes at any order on the ray, we can guarantee only that there exists a constant $c_k^1(s)>0$ s.t.

$$I_k^1(t) \le c_k^1(s) h^{d/2+2k}, \forall N \le k \le N+1.$$

Then, using Lemma 3.6.2, we have

$$\begin{aligned}\frac{\|\Box_h \overrightarrow{\phi}^h(t)\|^2_{\ell^2(h\mathbb{Z}^d)}}{E_h(\overrightarrow{\phi}^h(t),\partial_t\overrightarrow{\phi}^h(t))} &\le (N+3)\sum_{k=-1}^{N-1}\frac{c_k^1(s)}{C^-(s)}h^{2k+\beta_k+3} + (N+3)\sum_{k=N}^{N+1}\frac{c_k^1(s)}{C^-(s)}h^{2(k+1)} \\ &+ (N+1)\sum_{k=-2}^{N-2}\frac{c_k^2(s)}{C^-(s)}h^{2k+R+3}.\end{aligned} \tag{3.82}$$

The smallest order in the last term in (3.82) is h^{R-1}. The optimal β_k is obtained when one requires $2k+\beta_k+3 = R-1$. From $\beta_k \ge 0$ for all $-1 \le k \le N-1$, one obtains $R-2N \ge 2$. Moreover, $R-1 < 2(N+1)$, the lowest order in the second term in (3.82), otherwise one could solve the Eikonal equation up to order $R-1$ without loosing any order in the approximation of the ℓ^2-norm of $\Box_h \overrightarrow{\phi}^h(t)$. □

Lemma 3.6.4 (Estimate of the discrete energy outside a cylinder around the ray)**.** *For all $t>0$ and s s.t. $t = t(s)$, there exists a constant $C>0$ independent of h s.t., for all $r \in (h^{1/2},\delta)$, $\delta = (2-T|\nabla\omega_{d,1}(\eta_0)|)/2$ and ϕ the solution given by the ansatz (3.74), the following estimate holds:*

$$E_h(\overrightarrow{\phi}^h(\cdot,t(s)),\partial_t\overrightarrow{\phi}^h(\cdot,t(s))) \ge C\exp\big(\frac{r^2}{2h}\nu^*(s)\big)h^d \sum_{|x_{\mathbf{j}}-x(s)|\ge r}|\partial_t\phi(x_{\mathbf{j}},t(s))|^2, \tag{3.83}$$

with $\nu^*(s)$ the smallest eigenvalue of the matrix $N(s) = Im\big(D_x^2\psi(x(s),t(s))\big)$.

Remark 3.6.1. *The lower bound for r is obtained by considering $r=h^\beta$ and imposing $2\beta-1\le 0$, in order to ensure $r^2/2h \to \infty$ as $h\to 0$. The upper bounds is obtained by imposing $\{x : |x-x(s)| \le r\} \times \{s : t(s) \in (0,T)\} \subseteq B^d(0,1)\times(0,T)$. By considering $r = h^{1/4}$, the same divergence rate for the normalized energy in $|x-x(s)|>r$ of the Gaussian beam as in [42] is obtained. Nevertheless, observe that there exists the freedom to obtain a divergence rate $\exp(1/h)$ for the energy concentrated in $\Omega = \mathbb{R}^d \setminus B^d(0,1)$ of the Gaussian beam. But the Gaussian beam is an asymptotic solution, and in order to obtain the divergence rate of the observability constant for the finite-difference semi-discretization with the same initial data as those of the Gaussian beam, a perturbation technique has to be applied, which is the topic of the following lemma.*

Theorem 3.6.1. *Consider the equation (3.3) with the initial data given by*

$$(\overrightarrow{\phi}^{h,0}, \overrightarrow{\phi}^{h,1}) = (\phi(x_{\mathbf{j}}, 0), \partial_t \overrightarrow{\phi}(x_{\mathbf{j}}, 0))_{\mathbf{j}\in\mathbb{Z}^d},$$

with $\phi(x,t)$ given in the ansatz (3.74). Then there exists a constant $C > 0$ independent of h s.t.

$$\frac{E_h(\overrightarrow{\phi}^{h,0}, \overrightarrow{\phi}^{h,1})}{h^d \sum\limits_{x_{\mathbf{j}}\in\Omega} \int\limits_0^T |\partial_t \phi_{\mathbf{j}}(t)|^2 \, dt} \geq \frac{C}{\exp(-\delta^2\nu^*/h) + h^{R-1}},$$

where R is the order up to which the Eikonal $\widetilde{r}(x,t) = 0$ is solved on the ray $(x(s), t(s))$ and $\nu* = \inf\limits_{t(s)\in[0,T]} \nu^*(s)$.

Proof of Lemma 3.6.1. Considered the following corrector problem

$$\begin{cases} \Box_h v_{\mathbf{j}}(t) = -\Box_h \phi(x_{\mathbf{j}}, t), & \mathbf{j} \in \mathbb{Z}^d, t \in [0,T], \\ v_{\mathbf{j}}(0) = \partial_t v_{\mathbf{j}}(t) = 0, & \mathbf{j} \in \mathbb{Z}^d. \end{cases} \tag{3.84}$$

By multiplying in (3.84) by $\partial_t v_{\mathbf{j}}$ and taking sum in $\mathbf{j} \in \mathbb{Z}^d$, we obtain

$$E_h(\overrightarrow{v}^h(t), \partial_t \overrightarrow{v}^h(t)) \leq \|\Box_h \phi(\cdot, t)\|^2_{\ell^2(h\mathbb{Z}^d)}.$$

Observe that $\phi_{\mathbf{j}}(t) = \phi(x_{\mathbf{j}}, t) + v_{\mathbf{j}}(t)$. Then

$$h^d \sum_{x_{\mathbf{j}}\in\Omega} \int_0^T |\partial_t \phi_{\mathbf{j}}(t)|^2 \, dt \leq 2h^d \int_0^T \sum_{|x_{\mathbf{j}} - x(s)| > r} |\partial_t \phi(x_{\mathbf{j}}, t)|^2 \, dt + 2 \int_0^T \|\Box_h \overrightarrow{\phi}^h(\cdot, t)\|^2_{\ell^2(h\mathbb{Z}^d)} \, dt. \tag{3.85}$$

There exists a constant $C(s) > 0$ continuous in s and independent of h s.t.

$$E_h(\overrightarrow{\phi}^h(\cdot, t(s)), \partial_t \overrightarrow{\phi}^h(\cdot, t(s))) \sim C(s) E_h(\overrightarrow{\phi}^h(\cdot, 0), \partial_t \overrightarrow{\phi}^h(\cdot, 0)). \tag{3.86}$$

The inequality (3.84) follows from the above observations (3.85), (3.86) and by applying Lemmas 3.6.4 and 3.6.3. □

Remark 3.6.2. *It is important to observe that T cannot be considered to be infinite in the observability inequality. This corresponds to the fact that although the matrix $N(t) = Im(M(t))$ positive definite in d directions, as $t \to \infty$, it could degenerate in some of these directions. This holds even for the continuous wave equation. An example of this type is provided by J. Ralston in [47] for the $2-d$ wave equation. For the sake of completeness, we recall that example here: the ray of Geometric Optics is given by $(t, x(t), y(t)) = (t, 0, t)$ and the corresponding amplitude is*

$$\psi(t,x,y) = \frac{y-t}{2} + \frac{a^2x^2t}{1+4a^2t^2} + i\Big(\frac{a}{1+4a^2t^2}\frac{x^2}{2} + b\frac{(y-t)^2}{2}\Big),$$

with a, b two real parameters. The corresponding matrix $N(t)$ is given by

$$N(t) = \begin{pmatrix} b & 0 & -b \\ 0 & \frac{1}{1+4a^2t^2} & 0 \\ -b & 0 & b \end{pmatrix}.$$

The first line and columns of $N(t)$ correspond to the time variable. Remark that $N(t)$ is degenerate in this direction for all t. The matrix corresponding to the space variables has determinant

$$\frac{b}{1+4a^2t^2},$$

which for any finite time t is strictly positive, but vanishes as $t \to infty$. Therefore, as $t \to \infty$, $N(t)$ also degenerates in the x variable.

Chapter 4

Discontinuous Galerkin semi-discretizations of the $1-d$ wave equation

4.1 Introduction

In this chapter, we describe the propagation properties of two classes of discontinuous Galerkin (DG) approximations of the $1-d$ wave equation. More precisely, we deal with the so-called *Symmetric Interior Penalty Discontinuous Galerkin* (SIPG) (cf.[5]) and the *Local Discontinuous Galerkin* (LDG) (cf. [22]) methods. This topic is motivated, in particular, by control theoretical problems. By means of the Hilbert Uniqueness Method (HUM) introduced in [36], control problems for the wave equation are equivalent to appropriate observability inequalities for the adjoint system, consisting on determining whether the total energy of solutions of PDEs or of their numerical approximations can be estimated in terms of the energy concentrated on some subset of the spatial domain.

1. **The continuous $1-d$ wave equation.** Consider the following Cauchy problem associated to the $1-d$ wave equation:

$$\begin{cases} \partial_t^2 u(x,t) - \partial_x^2 u(x,t) = 0, & x \in \mathbb{R},\ t > 0 \\ u(x,0) = u^0(x),\ \partial_t u(x,0) = u^1(x), & x \in \mathbb{R}. \end{cases} \tag{4.1}$$

This problem is well posed for $(u^0, u^1) \in \dot{H}^1(\mathbb{R}) \times L^2(\mathbb{R})$. Namely, for any $(u^0, u^1) \in \dot{H}^1(\mathbb{R}) \times L^2(\mathbb{R})$, using the Hille-Yosida Theorem, $u \in C(\mathbb{R}_+, \dot{H}^1(\mathbb{R})) \cap C^1(\mathbb{R}_+, L^2(\mathbb{R}))$ and the total energy

$$E(u^0, u^1) = \frac{1}{2}\big(\|u(\cdot,t)\|^2_{\dot{H}^1(\mathbb{R})} + \|\partial_t u(\cdot,t)\|^2_{L^2(\mathbb{R})}\big) = \frac{1}{2}\big(\|u^0\|^2_{\dot{H}^1(\mathbb{R})} + \|u^1\|^2_{L^2(\mathbb{R})}\big), \tag{4.2}$$

is conserved in time. Here, $\dot{H}^1(\mathbb{R})$ denotes the complation of $C_c^\infty(\mathbb{R})$ with respect to the semi-norm $\|\cdot\|_{\dot{H}^1(\mathbb{R})} = \|\partial_x \cdot\|_{L^2(\mathbb{R})}$.

The **observability problem** (for both continuous and discrete cases) consists on determining whether the total energy of solutions can be estimated in terms of the energy concentrated on some subset of the spatial domain where waves propagate, the so-called *observation region*. For the continuous problem (4.1), it is well known that the observability property holds when the observation region is the complement of a compact set. More precisely, for any initial data $(u^0, u^1) \in \dot{H}^1(\mathbb{R}) \times L^2(\mathbb{R})$ and any observability time $T > T^* := 2$, there exists a constant $C(T) > 0$

s.t. the following *observability inequality* holds:

$$\boxed{E(u^0,u^1) \leqslant C(T) \int_0^T E_\Omega(u^0,u^1,t)\,dt,} \tag{4.3}$$

where $\Omega := \mathbb{R}\setminus(-1,1)$ and $E_\Omega(u^0,u^1,t)$ is called *the energy concentrated in* Ω associated to the initial data (u^0,u^1) at time t, given explicitly by

$$E_\Omega(u^0,u^1,t) = \frac{1}{2}\int_\Omega (|\partial_t u(x,t)|^2 + |\partial_x u(x,t)|^2)\,dx. \tag{4.4}$$

For simplicity, we assume $\Omega = \mathbb{R}\setminus(-1,1)$, but similar results hold for any $1-d$ exterior domain $\Omega = \mathbb{R}\setminus(a,b)$, where (a,b) is any finite interval, provided that $T > T^* := b-a$, and, as we have seen in the first chapter, also in several space dimensions. The best constant $C(T)$ in (4.3) is referred to as the *observability constant*.

The time $T^* := 2$ is sharp, given by the so-called *Geometric Control Condition* (GCC) (cf.[7]). This requires *all non-diffractive rays of Geometric Optics to touch the observation region during the observability time*. In the $1-d$ setting, this condition is particularly easy to see by means of the D'Alembert formula. More precisely, $T^* = 2$ is the time needed by a wave packet supported in an arbitrarily narrow neighborhood of one of the endpoints of the interval $(-1,1)$ at time $t=0$ to reach the other endpoint, traveling along the characteristics $x(t) = x \pm t$. The GCC is also valid in several space dimensions. Its proof uses microlocal tools and is inspired on the fact that the energy of the solutions propagates along the bi-characteristic rays ([7],[57]).

When the GCC does not hold, the observability property fails because one can build Gaussian beams localized around a bi-characteristic ray that escapes the observation region (see, for instance, [35],[37], [38], [46], [48] and [56]). In the $1-d$ case, this construction is even simpler, based on the D'Alembert formula. It consist on initial data of compact support of total energy one, such that the support of the solution at any time $t \in (0,T)$ do not intersect Ω. In this way, $E_\Omega(u^0,u^1,t) = 0$, for all $t \in (0,T)$, meaning that, in this particular $1-d$ setting, $C(T)$ in (4.3) may diverge at any order if the GCC fails.

This observability property (4.3) is motivated by controllability problems. More precisely, by means of HUM, the observability problem for system (4.1), called the adjoint system, is equivalent to the following *controllability problem*. For all $T > T^*$ and all $(y^0,y^1) \in \dot{H}^1(\mathbb{R}) \times L^2(\mathbb{R})$, there exists a control function $f \in L^\infty(0,T;L^2(\Omega))$ such that the solution of the following non-homogeneous problem

$$\begin{cases} \partial_t^2 y(x,t) - \partial_x^2 y(x,t) = f(x,t)\chi_\Omega(x), & x \in \mathbb{R},\ t \in (0,T] \\ y(x,0) = y^0(x),\ \partial_t y(x,0) = y^1(x),\ x \in \mathbb{R} \end{cases}$$

satisfies the equilibrium condition at time T:

$$y(x,T) = \partial_t y(x,T) = 0,\ \forall x \in \mathbb{R}.$$

These issues concerning the continuous wave equation are by now well understood and have been the object of intensive research. We refer to [60] for a recent survey on this and many other closely related topics.

2. **The finite difference semi-discretization of the $1-d$ wave equation.** As it has been recently observed, these issues are even more subtle when the continuous wave equation is replaced by a numerical scheme. For instance, it is by now well known that discrete analogous of the observability property may fail because of the pathological behavior of spurious high frequency numerical solutions, even if the continuous counterpart holds. This occurs, for instance, for the classical finite difference and finite element discretization schemes. We refer to [59] for a recent survey on this topic.

In order to make clear well known results concerning the propagation properties of some classical numerical approximations of the wave equation, we consider a mesh size $h > 0$, $x_j = jh$, $j \in \mathbb{Z}$, an uniform grid of the real line and introduce the Cauchy problem associated to the finite difference space semi-discretization of the $1-d$ wave equation:

$$\begin{cases} \partial_t^2 u_j(t) - \frac{u_{j+1}(t)-2u_j(t)+u_{j-1}(t)}{h^2} = 0, & j \in \mathbb{Z},\ t > 0 \\ u_j(0) = u_j^0,\ \partial_t u_j(0) = u_j^1, & j \in \mathbb{Z}. \end{cases} \tag{4.5}$$

This problem is well-posed in $\dot{h}^1(h\mathbb{Z}) \times \ell^2(h\mathbb{Z})$, where

$$\ell^2(h\mathbb{Z}) = \{ \overrightarrow{f} = (f_j)_{j\in\mathbb{Z}} \text{ s.t. } \| \overrightarrow{f} \|^2_{\ell^2(h\mathbb{Z})} := h \sum_{j\in\mathbb{Z}} |f_j|^2 < \infty \}$$

and

$$\dot{h}^1(h\mathbb{Z}) = \{ \overrightarrow{f} = (f_j)_{j\in\mathbb{Z}} \text{ s.t. } \| \overrightarrow{f} \|^2_{\dot{h}^1(h\mathbb{Z})} := \| \partial_h^+ \overrightarrow{f} \|^2_{\ell^2(h\mathbb{Z})} < \infty \},$$

with $\partial_h^+ := (f_{j+1} - f_j)/h$. Thus, for all initial data $(\overrightarrow{u}^{h,0}, \overrightarrow{u}^{h,1}) \in \dot{h}^1(h\mathbb{Z}) \times \ell^2(h\mathbb{Z})$ in the discrete problem (4.5), there exists an unique solution $\overrightarrow{u}^h(t) \in C^0(\mathbb{R}_+; \dot{h}^1(h\mathbb{Z})) \cap C^1(\mathbb{R}_+; \ell^2(h\mathbb{Z}))$. Moreover, its total energy is conserved in time and explicitly given by:

$$E_h(\overrightarrow{u}^{h,0}, \overrightarrow{u}^{h,1}) = \frac{1}{2}\big(\| \overrightarrow{u}^h(t) \|^2_{\dot{h}^1(h\mathbb{Z})} + \| \partial_t \overrightarrow{u}^h(t) \|^2_{\ell^2(h\mathbb{Z})} \big) = \frac{1}{2}\big(\| \overrightarrow{u}^{h,0} \|^2_{\dot{h}^1(h\mathbb{Z})} + \| \overrightarrow{u}^{h,1} \|^2_{\ell^2(h\mathbb{Z})} \big). \tag{4.6}$$

Given a time $T > 0$, consider the best constant $C_h(T) > 0$ such that for all $\overrightarrow{u}^{h,0} \times \overrightarrow{u}^{h,1} \in \dot{h}^1(h\mathbb{Z}) \times \ell^2(h\mathbb{Z})$ the following discrete version of the observability inequality (4.3) holds:

$$\boxed{E_h(\overrightarrow{u}^{h,0}, \overrightarrow{u}^{h,1}) \le C_h(T) \int_0^T E_{\Omega,h}(\overrightarrow{u}^{h,0}, \overrightarrow{u}^{h,1}, t)\, dt,} \tag{4.7}$$

with

$$E_{\Omega,h}(\overrightarrow{u}^{h,0}, \overrightarrow{u}^{h,1}, t) = \frac{1}{2}\big(\| \overrightarrow{u}^h(t) \|^2_{\dot{h}^1(\Omega)} + \| \partial_t \overrightarrow{u}^h(t) \|^2_{\ell^2(\Omega)} \big).$$

Although (4.6) is a convergent scheme for the continuous problem (4.1) in the classical sense of the numerical analysis, essential differences appear between the two observability inequalities (4.3) and (**??**). The most important is that for any finite time T, even when the GCC for the continuous wave equation holds, the observability constant $C_h(T)$ in (**??**) blows-up as h tends to zero. In the first chapter, using a result of Micu in [42] based on estimates on bi-orthogonal sequences, we proved that there exist initial data $(\overrightarrow{u}^{h,0}, \overrightarrow{u}^{h,1}) \in \dot{h}^1(h\mathbb{Z}) \times \ell^2(h\mathbb{Z})$ for which the observability inequality $C_h(T)$ diverges at least as $\exp(\sqrt{1/h})$ as $h \to 0$. We also analyzed the blow-up rate of the observability constant $C_h(T)$ for some examples of initial data in (4.5), for which we proved a polynomial divergence at any order.

By applying the SDFT at scale h (definition (4.4), Chapter 1), system (4.5) can be transformed into the following second-order ODE depending on the parameter ξ:

$$\begin{cases} \partial_t^2 \widehat{u}^h(\xi,t) + \Omega_h(\xi)\widehat{u}^h(\xi,t) = 0, & \xi \in \Pi_h, t \in \mathbb{R}_+ \\ \widehat{u}^h(\xi,0) = \widehat{u}^{h,0}(\xi), \quad \partial_t \widehat{u}^h(\xi,0) = \widehat{u}^{h,1}(\xi) & \xi \in \Pi_h, \end{cases} \tag{4.8}$$

whose solution is given by

$$\widehat{u}^h(\xi,t) = \sum_{\pm} \frac{1}{2}\Big(\widehat{u}^{h,0}(\xi) \pm \frac{\widehat{u}^{h,1}(\xi)}{i\omega_h(\xi)} \Big) \exp(\pm it\omega_h(\xi)), \tag{4.9}$$

where

$$\Omega_h(\xi) = \frac{4}{h^2} \sin^2\Big(\frac{\xi h}{2}\Big), \quad \omega_h(\xi) = \text{sign}(\xi)\sqrt{\Omega_h(\xi)}, \forall \xi \in \Pi_h. \tag{4.10}$$

In (4.8), $\Omega_h(\xi)$ is the Fourier symbol of the discrete Laplacian given by the three point scheme. For a fixed $\xi \in \mathbb{R}$, $\Omega_h(\xi) \to |\xi|^2$ as $h \to 0$, the limit being the Fourier symbol of the continuous Laplacian. This is due to the convergence of the approximation. The square root of $\Omega_h(\xi)$, $\omega_h(\xi)$, $\xi \in \Pi_h$, yields the frequencies involved in the solution of the semi-discrete wave equation, (4.9). The frequencies involved in the solution of the continuous wave equation (4.1) are $\omega(\xi) = \xi$, $\xi \in \mathbb{R}$. Observe that $\omega'(\xi) = 1$, for all $\xi \in \mathbb{R}$, whereas $\omega_h'(\xi) = \cos(\xi h/2)$. We see that $\omega_h'(\xi)$ is a strictly decreasing function on $(0, \pi/h)$, with $\omega_h'(0) = 1$ and $\omega_h'(\pi/h) = 0$. These first order derivatives, $\omega'(\xi)$ and $\omega_h'(\xi)$, are the so-called group velocities. In [54], the existence of wave packets propagating with the group velocity for some fully-discrete finite-difference schemes for the transport equation was showed. The amplitude of the Fourier transform of the wave packet is a rapidly decaying function at infinity, centered at some wave number ξ_0. This kind of solutions propagate along the so-called rays of Geometric Optics, which, roughly speaking, in the continuous case are $x(t) = x \pm t$, whereas in the finite-difference case they are $x_h(t) = x \pm \omega_h'(\xi_0)t$. The lack of a strictly positive lower bound for discrete group velocity $\omega_h'(\xi)$ means that there are rays arbitrarily close to the vertical position or wave packets propagating very slow, which, in order to travel a finite distance spent a very large time.

3. The SIPG semi-discretization of the $1-d$ wave equation. Based on discontinuous finite-element spaces, the discontinuous Galerkin (DG) methods easily handle elements of various types and shapes, irregular non-matching grids and even varying polynomial order. They were proposed in the seventies for the numerical solution of hyperbolic neutron transport equations (cf. [28]) and, independently, for elliptic and parabolic problems (cf. [6]). Usually, the DG methods for the two last classes of problems are called *interior penalty* (IP) methods. In the IP methods, the continuity is weakly enforced across the element interfaces, by adding suitable bilinear forms called numerical fluxes to the classical variational formulations.

In recent years, an intensive research has been developed on DG methods. We refer to [6], [13], [14] for an unified analysis and a comparison of all the existing DG methods proposed for the numerical treatment of the elliptic problems. In [4], the problem of computing eigenvalues and eigenfunctions of the Laplace operator by means of DG methods is analyzed and and it is showed that several DG methods provide a spectrally correct approximation of the Laplace operator. The hermitian DG methods provide optimal approximation of the eigenfunctions and the eigenvalues of the Laplace operator. In the context of the wave equation, we refer to [1], where the dispersive and dissipative properties of hp version of DG methods are studied in several limits. For small wave number $\xi h \to 0$, it was showed that the DG methods provide a higher order of accuracy than the classical Galerkin methods. By keeping the mesh size h fixed and increasing the polynomial order p, the dissipation and dispersion errors are shown to decay at a super-exponential rate when the order p is much larger than ξh and exponentially if p and ξh are of the same order. In [2], explicit lower bounds for the stability parameter in the $1-d$ SIPG method are obtained in order to avoid pollution by spurious modes. In [27], optimal a priori error bounds are obtained in the energy norm and in the L^2-norm for the SIPG semi-discretization of the multi-dimensional wave equation.

In this section, we deal with the simplest setting of the SIPG space semi-discretization of the $1-d$ wave equation: an uniform grid x_j, $j \in \mathbb{Z}$, and first order polynomials.

The two important quantities in a DG method are the average and the jump of the numerical solution along the interface. In the $1-d$ setting, we define the jump $[\cdot]$ and the average $\{\cdot\}$ of a function f at the point x to be

$$[f](x) = f(x-) - f(x+), \quad \{f\}(x) = \frac{f(x+) + f(x-)}{2}.$$

Similar definitions can be given in the multi-dimensional case (cf. [6]). We introduce the broken Sobolev space

$$H_h^1(\mathbb{R}) = \{f \in L^2(\mathbb{R}) : \|f\|_h^2 := \sum_{j\in\mathbb{Z}} \int_{x_j}^{x_{j+1}} |f_x(x)|^2\, dx + \frac{1}{h}\sum_{j\in\mathbb{Z}} [f]^2(x_j) < \infty\}.$$

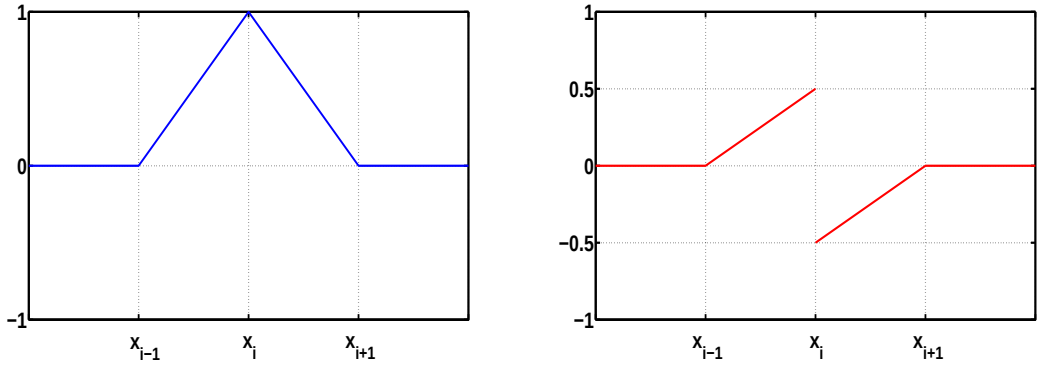

Figure 4.1: Typical basis functions for the P_1-DG methods: ϕ_i^A (left) and ϕ_i^J (right).

Set P_h to be the set of polynomials of degree at most 1 in each interval (x_j, x_{j+1}), $j \in \mathbb{Z}$.

The finite element space is given by $V_h = H_h^1(\mathbb{R}) \cap P_h$. Observe that $V_h = U_h^A \oplus U_h^J$, with $U_h^A = \text{span}\{\phi_i^A, i \in \mathbf{Z}\} \cap H^1(\mathbb{R})$ and $U_h^J = \text{span}\{\phi_i^J, i \in \mathbf{Z}\}$, where

$$\phi_i^A(x) = \Big[1 - \frac{|x - x_i|}{h}\Big]^+, \quad \phi_i^J(x) = \frac{1}{2}\text{sign}(x_i - x)\Big[1 - \frac{|x - x_i|}{h}\Big]^+.$$

Remark that ϕ_i^A is the typical basis function used in the P_1-classical finite element method, whereas ϕ_i^J, having average zero and jump one in x_i, is designed to represent the jump at the nodal point x_i. Each element $F \in V_h$ has an unique representation as a linear combination of the form

$$f(x) = \sum_{i \in \mathbf{Z}} f_i^A \phi_i^A(x) + \sum_{i \in \mathbf{Z}} f_j^J \phi_i^J(x) = f^A(x) + f^J(x),$$

where f^A and f^J are the continuous and the jump components of f, respectively. In this way, the piecewise linear discontinuous functions under consideration are perturbations of the classical piecewise linear and continuous ones by jumps added at each nodal point.

Before defining the bilinear form associated to the SIPG method, let us introduce the following auxiliary bilinear forms defined on $V_h \times V_h$:

$$b_h(u,v) = \sum_{j \in \mathbb{Z}} \int_{x_j}^{x_{j+1}} u_x(x) v_x(x)\, dx, \quad c_h(u,v) = -\sum_{j \in \mathbb{Z}} ([u](x_j)\{v_x\}(x_j) + [v](x_j)\{u_x\}(x_j)),$$
$$d_h(u,v) = \sum_{j \in \mathbb{Z}} [u](x_j)[v](x_j), \tag{4.11}$$

which generate block-tri-diagonal matrices generated by the stencils

$$r_{b_h} = \left(\begin{array}{cc|cc|cc} -\frac{1}{h} & \frac{1}{2h} & \frac{2}{h} & 0 & -\frac{1}{h} & -\frac{1}{2h} \\ -\frac{1}{2h} & \frac{1}{4h} & 0 & \frac{1}{2h} & \frac{1}{2h} & \frac{1}{4h} \end{array}\right), \quad r_{c_h} = \left(\begin{array}{cc|cc|cc} 0 & -\frac{1}{2h} & 0 & 0 & 0 & \frac{1}{2h} \\ \frac{1}{2h} & -\frac{1}{2h} & 0 & -\frac{1}{h} & -\frac{1}{2h} & -\frac{1}{2h} \end{array}\right)$$
$$r_{d_h} = \left(\begin{array}{cc|cc|cc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right). \tag{4.12}$$

For each $s > 1$, the so-called *penalty parameter*, we introduce the following symmetric bilinear form $a_h^s : V_h \times V_h \to \mathbb{R}$, given by

$$a_h^s(u,v) = b_h(u,v) + c_h(u,v) + \frac{s}{h} d_h(u,v). \tag{4.13}$$

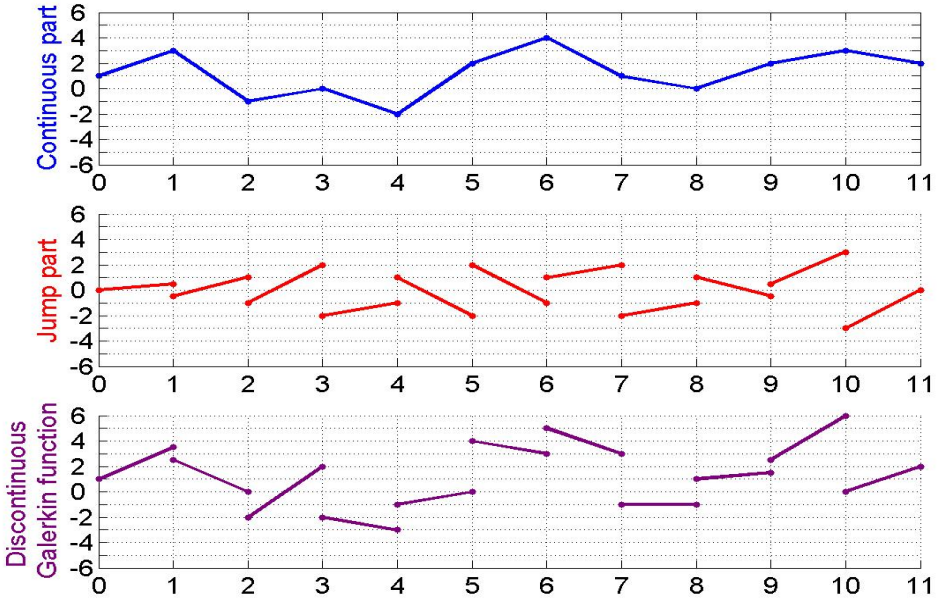

Figure 4.2: Example of a DG function and its decomposition into its continuous and jump parts.

Using this bilinear form in the approximation of the Laplacian in the continuous wave equation (4.1), we consider the following variational formulation of the SIPG semi-discrete wave equation:

$$\boxed{\text{Find } u_h^s(\cdot,t)\in V_h,\ \text{s.t. } \partial_t^2(u_h^s(\cdot,t),v)_{L^2(\mathbb{R})}+a_h^s(u_h^s(\cdot,t),v)=0,\ \forall v\in V_h,\ \forall t\geq 0,} \tag{4.14}$$

complemented with initial data $u_h^s(\cdot,0)=u_h^0\in V_h$ and $\partial_t u_h^s(\cdot,0)=u_h^1\in V_h$.

The unknown $u_h^s(x,t)$, being an element of V_h for each $t>0$, has the following decomposition

$$u_h^s(x,t)=\sum_{k\in\mathbf{Z}}A_k(t)\phi_k^A(x)+\sum_{k\in\mathbf{Z}}J_k(t)\phi_k^J(x).$$

Using all the functions ϕ_k^A and ϕ_k^J as test functions in (4.14), we obtain that the sequences of coefficients $\overrightarrow{U}^h(t)=(A_k(t),J_k(t))_{k\in\mathbb{Z}}$ of the approximation u_h^s are the solutions of the following infinite system of second-order differential equations:

$$\begin{cases} m_h\left(\partial_t^2A_{k-1}(t)\ \partial_t^2J_{k-1}(t)\ \partial_t^2A_k(t)\ \partial_t^2J_k(t)\ \partial_t^2A_{k+1}(t)\ \partial_t^2J_{k+1}(t)\right)' \\ \qquad +r_h^s\left(A_{k-1}(t)\ J_{k-1}(t)\ A_k(t)\ J_k(t)\ A_{k+1}(t)\ J_{k+1}(t)\right)'=(0\ 0)\,,\ k\in\mathbb{Z},t>0, \\ \overrightarrow{A}^h(0)=\overrightarrow{A}^{h,0},\ \partial_t\overrightarrow{A}^h(0)=\overrightarrow{A}^{h,1},\ \overrightarrow{J}^h(0)=\overrightarrow{J}^{h,0},\partial_t\overrightarrow{J}^h(0)=\overrightarrow{J}^{h,1}, \end{cases} \tag{4.15}$$

where m_h, r_h^s are the matrices

$$m_h=\begin{pmatrix} \frac{h}{6} & -\frac{h}{12} & \frac{2h}{3} & 0 & \frac{h}{6} & \frac{h}{12} \\ \frac{h}{12} & -\frac{h}{24} & 0 & \frac{h}{6} & -\frac{h}{12} & -\frac{h}{24} \end{pmatrix},\ r_h^s=\begin{pmatrix} -\frac{1}{h} & 0 & \frac{2}{h} & 0 & -\frac{1}{h} & 0 \\ 0 & -\frac{1}{4h} & 0 & \frac{2s-1}{2h} & 0 & -\frac{1}{4h} \end{pmatrix}.$$

The system (4.15) can be also written in the following matrix form:

$$\begin{cases} M_h\partial_t^2\overrightarrow{U}^h(t)+R_h^s\overrightarrow{U}^h(t)=0, & t>0 \\ \overrightarrow{U}^h(0)=\overrightarrow{U}^{h,0}, \quad \partial_t\overrightarrow{U}^h(0)=\overrightarrow{U}^{h,1}, \end{cases} \tag{4.16}$$

where M_h and R_h^s are the mass and the stiffness matrices obtained as an infinite repetition of the stencils m_h, r_h^s. Both of them are block tri-diagonal and symmetric.

We can do several suitable choices for the initial data in (4.15). One of them is when the sequences $\overrightarrow{A}^{h,i}$ and $\overrightarrow{J}^{h,i}$, $i = 0, 1$, are L^2-projections of the initial data $(u^0, u^1) \in \dot{H}^1(\mathbb{R}) \times L^2(\mathbb{R})$ corresponding to the continuous wave equation (4.1) on the subspaces U_h^A and U_h^J, i.e., for $i = 0, 1$, $\overrightarrow{U}^{h,i} = (\overrightarrow{A}^{h,i}, \overrightarrow{J}^{h,i})$ is the solution of the system

$$M_h \overrightarrow{U}^{h,i} = \overrightarrow{F}^{h,i}, \tag{4.17}$$

where $\overrightarrow{F}^{h,i} = (F_k^i)_{k\in\mathbb{Z}}$, with

$$F_k^i = \Big(\int_{\mathbb{R}} u^i(x)\phi_k^A(x)\,dx \quad \int_{\mathbb{R}} u^i(x)\phi_k^J(x)\,dx \Big).$$

Another one is $(A_k^{h,i}, J_k^{h,i}) = (u^i(x_k), 0)$, which requires u^i to be continuous at the grid points. When the initial data for the continuous problem u^i are not continuous at the grid points x_j, we consider $A_k^{h,i}$ to be the average of u^i on the cell $[x_{k-1/2}, x_{k+1/2}]$.

The aim of this chapter is to analyze the SIPG version of the observability inequality (4.3). More precisely, for $\Omega = \mathbb{R} \setminus (-1, 1)$ as before, we investigate under which conditions on the observability time $T > 0$ and on the initial data $\overrightarrow{U}^{h,i}$ in (4.15), $i = 0, 1$, the following observability inequality holds

$$\boxed{E_h^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}) \le C_h^s(T) \int_0^T E_{\Omega,h}^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\,dt,} \tag{4.18}$$

where the discrete total energy is conserved in time and and defined as

$$E_h^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}) = \frac{1}{2}\Big(\langle R_h^s \overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,0}\rangle + \langle M_h \overrightarrow{U}^{h,1}, \overrightarrow{U}^{h,1}\rangle\Big).$$

By $\langle\cdot,\cdot\rangle$, we denote the inner product in $\ell^2(\mathbb{Z})$. The energy concentrated on Ω is given by

$$E_{\Omega,h}^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t) = \frac{1}{2}\Big(\langle R_h^s \overrightarrow{U}^h(t), \overrightarrow{U}^h(t)\rangle_\Omega + \langle M_h \partial_t \overrightarrow{U}^h(t), \partial_t \overrightarrow{U}^h(t)\rangle_\Omega\Big),$$

where $\langle\cdot,\cdot\rangle_\Omega$ denotes the inner product in $\ell^2(\{j : x_j \in \Omega\})$. In (4.27) and (4.26) we give a more precise definition of these energies, better adapted to the ℓ^2-setting and to use the discrete Parseval identity (3.14).

In what follows, we briefly describe the results obtained in this chapter.

Section 4.2 deals with the Fourier analysis of the Cauchy problem (4.5) for the finite difference semi-discretization viewed as a system in which odd and even components interact. This system produces two dispersion diagrams: one of them representing the low frequencies and the other one representing the high frequencies. In fact, the first one is ω_h (introduced in (4.10)) restricted to $[0, \pi/2h]$ and the second one is ω_h restricted to $[\pi/2h, \pi/h]$ and reflected with respect to $\pi/2h$. The interesting fact is that the group velocity corresponding to the high frequency-type diagram vanishes at the wave number $\xi = 0$, a phenomena which appears for the DG semi-discretization on the spurious diagram very often.

Section 4.3 is concerned to a careful Fourier analysis of the system (4.15). After a rearrangement of the terms in the total energy better adapted to apply the Parseval identity in Subsection 4.3.1, in Subsection 4.3.2 we we deal with the Fourier counterpart of (4.15). The Fourier analysis reveals two dispersion diagrams: a *physical* one, which for small values of the penalty parameter s closed to 1 behaves like the finite difference dispersion relation and for large s like the one corresponding to the P_1-classical finite element method. Also there exists a *spurious* diagram, which for small values of s is similar to a finite difference diagram, but for large values of s tends

to infinity. The corresponding eigenvectors provide also a rich phenomenology, with an essential change of their behavior as the wave number $\xi \to \pi/h$ at $s = 3$. Concerning the corresponding *group velocities*, several ranges of s can be identified:

a) $s \in (5/2, \infty) \setminus \{3\}$. The spurious diagram is decreasing on $(0, \pi/h)$, with two singular points located at $\xi = 0$ and at $\xi = \pi/h$ and the physical dispersion is strictly increasing with a singular point at $\xi = \pi/h$. See Figure 4.5, bottom - right.

b) $s = 3$. The global shape of the two dispersion relations is very similar to what happens for the finite difference semi-discretization viewed as a system where odd and even components interact. The spurious diagram is decreasing, the physical one is increasing and only a singular point appears on the spurious dispersion relation at the wave number $\xi = 0$. See Figure 4.5, bottom - left.

c) $s \in (5/3, 5/2)$. The physical diagram is increasing with a singular point at $\xi = \pi/h$, and the spurious one has three critical points located at the wave numbers $\xi = 0$, $\xi = \pi/h$ and a maximum point in $(0, \pi/h)$. See Figure 4.5, top - right.

d) $s \in (1, 5/3)$. The physical dispersion relation is increasing, with a critical point at $\xi = \pi/h$, and the spurious one is also increasing, with two critical points at $\xi = 0$ and $\xi = \pi/h$. See Figure 4.5, top - left.

Several remedies for avoiding the high frequency spurious oscillations have been developed in the existing literature: Tychonoff regularization, multi-grid methods, mixed finite element methods, numerical viscosity and Fourier filtering of high frequencies. In this chapter we will use only two of these filtering algorithms: *Fourier truncation of the high frequencies*, used for instance in ([33]) in order to guarantee the uniform observability for the finite difference and the classical P_1-finite element space semi-discretizations of the $1-d$ wave equation. It consists on allowing initial data whose SDFTs are supported on $[-\delta\pi/h, \delta\pi/h]$ for some $\delta \in (0, 1)$. This is an algorithm efficient from a theoretical point of view, but difficult to implement. A more convenient strategy for the practical implementation is *the bi-grid algorithm* (see [45] for the $1-d$ case and [32] for the analysis of the method for the wave equation in the unit square), dealing with initial data in the fine grid obtained by linear interpolation from data given on a coarser grid.

In **Section 4.5**, we analyze the following four filtering algorithms, each one of them requiring to combine two techniques.

Algorithm A: *Initial data involving only the physical mode and obtained by a Fourier truncation* of parameter $\delta \in (0, 1)$, analyzed in Subsection 4.5.1. Our methodology of proof does not allow to obtain the optimal observability time indicated by the GCC. Since the main result in this subsection, Theorem 4.5.1, will be used within the analysis of the other algorithms, the observability times obtained from the other algorithms are not optimal.

Algorithm B: *Initial data involving only the physical mode and obtained by a bigrid algorithm* of ratio 1/2, analyzed in Subsection 4.5.2. The main result in this subsection, Theorem 4.5.2, is based on an estimate of the total energy in terms of the energy concentrated on the physical dispersion relation restricted to Π_{2h}, corresponding to the low frequencies. Also a dyadic decomposition argument and Theorem 4.5.1 for the particular case $\delta = 1/2$ are used.

Algorithm C: *Initial data with null jump part and obtained by Fourier truncation* of parameter $\delta \in (0, 1)$, analyzed in Subsection 4.5.3. The main result in this subsection, Theorem 4.5.3, is based on an estimate of the total energy in terms of the energy concentrated on the physical diagram.

Algorithm D: *Initial data with null jump part and obtained by a bigrid algorithm* of ratio 1/2, analyzed in Subsection 4.5.4. The main result in this subsection, Theorem 4.5.4, is based on two estimates: one of them, establishing an upper bound of the total energy in terms of the energy concentrated on both dispersion relations restricted to Π_{2h}, and the other one, an upper bound of the energy concentrated on both dispersion relations restricted to Π_{2h} in terms of the energy concentrated on the physical dispersion relation restricted to Π_{2h}.

4.2 The finite difference semi-discretization as a system

In this section, we show that the most classical semi-discretizations of the wave equation like the finite difference one can be seen as a complex system yielding two dispersion relations: one representing the low frequencies and another one representing the high frequencies. Moreover, the group velocity corresponding to the high frequencies dispersion relation vanishes as the wave number ξ tends to zero, a pathology that appear at the discontinuous Galerkin semi-discretization.

Consider the equivalent system form of (4.5):

$$\left\{\begin{array}{ll} \partial_t^2 E_j(t) - \frac{O_j(t)-2E_j(t)+O_{j-1}(t)}{h^2} = 0, & j\in\mathbb{Z},\ t>0 \\ \partial_t^2 O_j(t) - \frac{E_{j+1}(t)-2O_j(t)+E_j(t)}{h^2} = 0, & j\in\mathbb{Z},\ t>0 \\ E_j(0)=E_j^0,\ O_j(0)=O_j^0,\ \partial_t E_j(0)=E_j^1,\ \partial_t O_j(0)=O_j^1, & j\in\mathbb{Z}, \end{array}\right. \tag{4.19}$$

where $E_j(t)=u_{2j}(t)$ and $O_j(t)=u_{2j+1}(t)$.

Observe that $\overrightarrow{E}^{2h}(t)$ and $\overrightarrow{O}^{2h}(t)$ corresponds to a grid of size $2h$. Then their SDFTs are π/h-periodic. Moreover, for all $\xi\in\Pi_h$, we have the following identity relating the SDFT at scale h of $\overrightarrow{u}(t)$ and the SDFT's at the scale $2h$ of $\overrightarrow{E}^{2h}(t)$ and $\overrightarrow{O}^{2h}(t)$:

$$\widehat{u}^h(\xi,t)=\frac{1}{2}\big(\widehat{E}^{2h}(\xi,t)+\widehat{O}^{2h}(\xi,t)\exp(-i\xi h)\big). \tag{4.20}$$

Fourier analysis of the finite-difference system. Consider $\overrightarrow{U}^{2h}(\xi,t)=(\overrightarrow{E}^{2h}(\xi,t),\overrightarrow{O}^{2h}(\xi,t))'$. Then by applying the SDFT at scale $2h$ to the system (4.19), we have:

$$\left\{\begin{array}{ll} \partial_t^2\widehat{U}^{2h}(\xi,t)+A_h(\xi)\widehat{U}^{2h}(\xi,t)=0, & \xi\in\Pi_{2h}, t>0 \\ \widehat{U}^{2h}(\xi,0)=\widehat{U}^{2h,0}(\xi), \quad \partial_t\widehat{U}^{2h}(\xi,0)=\widehat{U}^{2h,1}(\xi) & \xi\in\Pi_{2h}, \end{array}\right. \tag{4.21}$$

where $A_h(\xi)$ is a matrix given by

$$A_h(\xi)=\begin{pmatrix} \frac{2}{h^2} & -\frac{1+\exp(-2i\xi h)}{h^2} \\ -\frac{1+\exp(2i\xi h)}{h^2} & \frac{2}{h^2} \end{pmatrix}. \tag{4.22}$$

The two eigenvalues of $A_h(\xi)$ are

$$\Lambda_h^{lo}(\xi)=\frac{4}{h^2}\sin^2\Big(\frac{\xi h}{2}\Big), \quad \Lambda_h^{hi}(\xi)=\frac{4}{h^2}\cos^2\Big(\frac{\xi h}{2}\Big) \tag{4.23}$$

and the corresponding eigenvectors are

$$P_h^{lo}(\xi)=\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}}\exp(i\xi h) \end{pmatrix}, \quad P_h^{hi}(\xi)=\begin{pmatrix} -\frac{1}{\sqrt{2}}\exp(-i\xi h) \\ \frac{1}{\sqrt{2}} \end{pmatrix}. \tag{4.24}$$

Denote by $\lambda_h^{lo}(\xi)$ and $\lambda_h^{hi}(\xi)$ the square roots of $\Lambda_h^{lo}(\xi)$ and $\Lambda_h^{hi}(\xi)$ for $\xi\in(0,\pi/2h)$ and the odd extensions for $\xi\in(-\pi/2h,0)$.

For all $\xi\in\Pi_{2h}$, the solution of (4.21) is given by

$$\begin{aligned}\widehat{U}^{2h}(\xi,t)=\sum_{\pm}\frac{1}{2}\Big[&P_h(\xi)\begin{pmatrix} \exp(\pm it\lambda_h^{lo}(\xi)) & 0 \\ 0 & \exp(\pm it\lambda_h^{hi}(\xi)) \end{pmatrix}\big(P_h(\xi)\big)^{-1}\widehat{U}^{2h,0}(\xi) \\ &+P_h(\xi)\begin{pmatrix} \pm\frac{\exp(\pm it\lambda_h^{lo}(\xi))}{i\lambda_h^{lo}(\xi)} & 0 \\ 0 & \pm\frac{\exp(\pm it\lambda_h^{hi}(\xi))}{i\lambda_h^{hi}(\xi)} \end{pmatrix}\big(P_h(\xi)\big)^{-1}\widehat{U}^{2h,1}(\xi)\Big],\end{aligned} \tag{4.25}$$

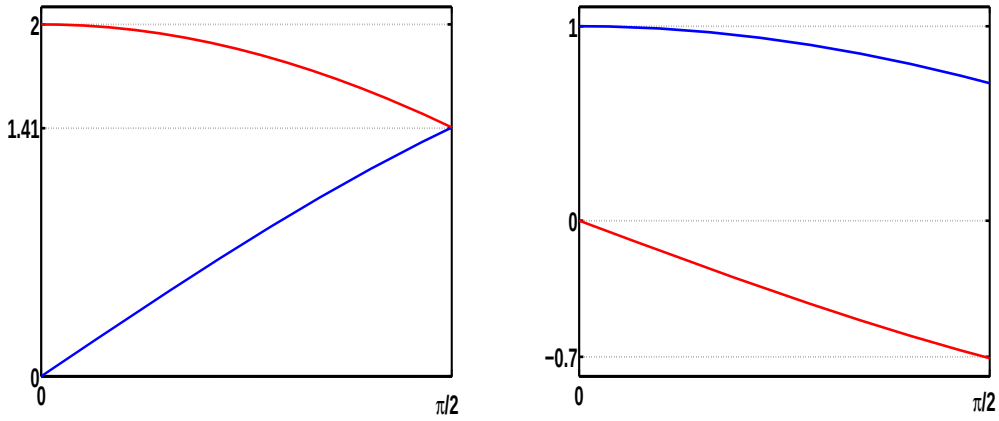

Figure 4.3: Dispersion relations $\lambda_1^{lo}(\xi)$ (blue) and $\lambda_1^{hi}(\xi)$ (red) (left) and the corresponding group velocities (right).

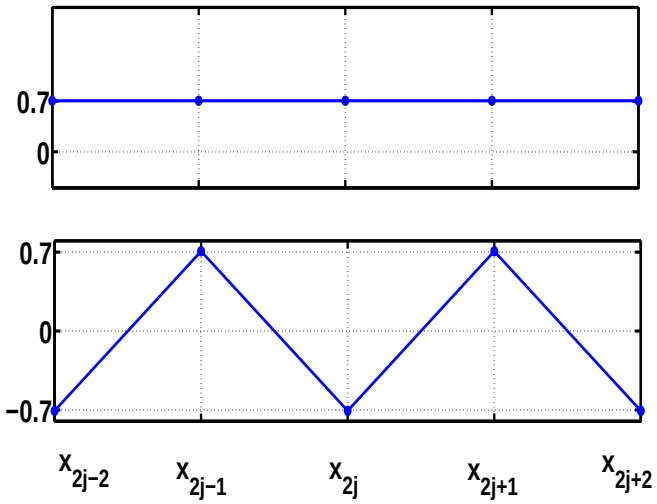

Figure 4.4: The low frequency eigenvector, $P_h^{lo}(\xi)$, versus the high frequency one, $P_h^{hi}(\xi)$, at the wave number $\xi = 0$.

where $P_h(\xi)$ is the matrix whose columns are the eigenvectors $P_h^{lo}(\xi)$ and $P_h^{hi}(\xi)$ and $(P_h(\xi))^{-1}$ is its inverse.

In Figure 4.4, we represent a discrete function $(F_j)_{j\in\mathbb{Z}} = (E_j, O_j)'_{j\in\mathbb{Z}}$ whose Fourier representation is $\widehat{F}^{2h}(\xi) = 2\pi P_h^{lo}(\xi)\delta_0(\xi)$ (top) and $\widehat{F}^{2h}(\xi) = 2\pi P_h^{hi}(\xi)\delta_0(\xi)$ (bottom).

Proposition 4.2.1. *The eigenvalues $\Lambda_h^{lo}(\xi)$, $\Lambda_h^{hi}(\xi)$ and the eigenvectors $P_h^{lo}(\xi)$ and $P_h^{hi}(\xi)$ have the following properties:*

FD1. $\lim_{\xi\to 0} \lambda_h^{lo}(\xi) = 0$, $\lim_{\xi\to 0} \lambda_h^{hi}(\xi) = 2/h$.

FD2. $\lim_{\xi\to 0} P_h^{lo}(\xi) = (1/\sqrt{2}, 1/\sqrt{2})'$, $\lim_{\xi\to 0} P_h^{hi}(\xi) = (-1/\sqrt{2}, 1/\sqrt{2})'$.

FD3. $\lim_{\xi\to 0}\partial_\xi\lambda_h^{lo}(\xi)=1$, $\lim_{\xi\to 0}\partial_\xi\lambda_h^{hi}(\xi)=0$ *and for all* $\xi\in(0,\pi/2h)$, $\partial_\xi\lambda_h^{lo}(\xi)=\cos(\xi h/2)>0$ *and* $\partial_\xi\lambda_h^{hi}(\xi)=-\sin(\xi h/2)<0$.

Then $\lambda_h^{lo}(\xi)$ is strictly increasing and $\lambda_h^{hi}(\xi)$ is strictly decreasing in $\xi\in[0,\pi/h]$.

4.3 The symmetric interior penalty space semi-discretization

4.3.1 The $\ell^2(h\mathbb{Z})$ form of the discrete energy

In what follows, we consider the SIPG semi-discretization (4.15) of the wave equation. In this subsection, we organize the total energy and the energy concentrated on Ω as a sum of $\ell^2(h\mathbb{Z})$ norms, this approach being more appropriate to find the Fourier representation of the energy by means of the Parseval identity and also in the proof of the first result concerning the filtering mechanisms, Theorem 4.5.1.

For $O\subset\mathbb{R}$, we define the energy concentrated in the set O at time t to be

$$E_{O,h}^s(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\sum_{k=1}^{7}E_{O,h}^{s,k}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t),\tag{4.26}$$

with

$$E_{O,h}^{s,1}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{h}{24}\sum_{x_k\in O}\left[|\partial_tA_{k+1}(t)+\partial_tA_k(t)|^2+|\partial_tA_k(t)+\partial_tA_{k-1}(t)|^2\right],$$

$$E_{O,h}^{s,2}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{h}{4}\sum_{x_k\in O}\left[\left|\frac{A_{k+1}(t)-A_k(t)}{h}\right|^2+\left|\frac{A_k(t)-A_{k-1}(t)}{h}\right|^2\right],$$

$$E_{O,h}^{s,3}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{h}{96}\sum_{x_k\in O}\left[|\partial_tJ_{k+1}(t)-\partial_tJ_k(t)|^2+|\partial_tJ_k(t)-\partial_tJ_{k-1}(t)|^2\right],$$

$$E_{O,h}^{s,4}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{h}{16}\sum_{x_k\in O}\left[\left|\frac{J_{k+1}(t)-J_k(t)}{h}\right|^2+\left|\frac{J_k(t)-J_{k-1}(t)}{h}\right|^2\right],$$

$$E_{O,h}^{s,5}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{s-1}{2h}\sum_{x_k\in O}|J_k(t)|^2,$$

$$E_{O,h}^{s,6}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{h}{24}\sum_{x_k\in O}\left[\left|\partial_tA_k(t)+\frac{1}{2}\partial_tJ_{k+1}(t)\right|^2+\left|\partial_tA_{k-1}(t)+\frac{1}{2}\partial_tJ_k(t)\right|^2\right],$$

$$E_{O,h}^{s,7}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{h}{24}\sum_{x_k\in O}\left[\left|\partial_tA_{k+1}(t)-\frac{1}{2}\partial_tJ_k(t)\right|^2+\left|\partial_tA_k(t)-\frac{1}{2}\partial_tJ_{k-1}(t)\right|^2\right].$$

Define the total energy associated to the solution of (4.15) to be

$$E_h^s(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t):=E_{\mathbb{R},h}^s(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t).\tag{4.27}$$

The following property concerning the conservation of the total energy holds:

Proposition 4.3.1. *The total energy associated to the solution of (4.15) is conserved in time, i.e. for all* $t\geq 0$ *we have*

$$E_h^s(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=E_h^s(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},0).\tag{4.28}$$

For this reason, in what follows, we skip the argument t by denoting $E_h^s(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1})$ the total energy of the solutions of (4.15) associated to the initial data $(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1})$.

Proof. Since the matrices M_h and R_h^s involved in (4.15) are symmetric, we can say from the very beginning that the total energy of the solution is conserved in time. But in order to see how the seven terms involved in the energy arise, we provide a proof by using multipliers. Thus, let us multiply the first equation in (4.15) by $\partial_t \overline{A}_k(t)$, its complex conjugate by $\partial_t A_k(t)$, add the identities so obtained and add the resulting identity in $k \in \mathbb{Z}$. We get

$$E_1 + \partial_t\Big(E_h^{s,1}(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t) + \frac{h}{6}\sum_{k\in\mathbb{Z}} |\partial_t A_k(t)|^2 + E_h^{s,2}(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\Big) = 0, \tag{4.29}$$

with

$$E_1 = \frac{h}{24}\sum_{k\in\mathbb{Z}} \big(\partial_t^2 J_{k+1}(t)\partial_t \overline{A}_k(t) - \partial_t \overline{A}_{k+1}(t)\partial_t^2 J_k(t) + \partial_t^2 \overline{J}_{k+1}(t)\partial_t A_k(t) - \partial_t A_{k+1}(t)\partial_t^2 \overline{J}_k(t)\big).$$

In the same way, we multiply the second equation in (4.15) by $\partial_t \overline{J}_k(t)$, its complex conjugate by $\partial_t J_k(t)$, adding the two identities so obtained and the resulting identity in $k \in \mathbb{Z}$, we get

$$E_2 + \partial_t\Big(E_h^{s,3}(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t) + E_h^{s,4}(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t) + E_h^{s,5}(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t) + \frac{h}{24}\sum_{k\in\mathbb{Z}} |\partial_t J_k(t)|^2\Big) = 0, \tag{4.30}$$

with

$$E_2 = \frac{h}{24}\sum_{k\in\mathbb{Z}} \big(\partial_t \overline{J}_{k+1}(t)\partial_t^2 A_k(t) - \partial_t^2 A_{k+1}(t)\partial_t \overline{J}_k(t) + \partial_t J_{k+1}(t)\partial_t^2 \overline{A}_k(t) - \partial_t^2 \overline{A}_{k+1}(t)\partial_t J_k(t)\big).$$

By adding the identities (4.29) and (4.30) and taking into account the fact that

$$E_1 + E_2 = \partial_t\Big(E_h^{s,6}(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t) + E_h^{s,7}(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t) - \tfrac{h}{6}\sum_{k\in\mathbb{Z}} |\partial_t A_k(t)|^2 - \tfrac{h}{24}\sum_{k\in\mathbb{Z}} |\partial_t J_k(t)|^2\Big),$$

we obtain that $\partial_t E_h^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t) = 0$, which concludes the conservation in time of the total energy. □

4.3.2 Fourier analysis of the SIPG method

For $\xi \in \Pi_h$, let us denote by $\widehat{A}^h(\xi, t)$, $\widehat{J}^h(\xi, t)$ the SDFT's of the sequences of averages, $\overrightarrow{A}^h(t)$, and jumps, $\overrightarrow{J}^h(t)$. Similarly, by $\widehat{A}^{h,i}(\xi)$, $\widehat{J}^{h,i}(\xi)$, $i = 0, 1$, we denote the SDFT's of the initial data $\overrightarrow{A}^{h,i}$, $\overrightarrow{J}^{h,i}$. Set $\widehat{U}^h(\xi, t) := (\widehat{A}^h(\xi, t), \widehat{J}^h(\xi, t))^T$.

The Fourier symbols of the mass and stiffness matrices are

$$M_h(\xi) = \begin{pmatrix} \frac{2+\cos(\xi h)}{3} & \frac{i\sin(\xi h)}{6} \\ -\frac{i\sin(\xi h)}{6} & \frac{2-\cos(\xi h)}{12} \end{pmatrix}, \quad R_h^s(\xi) = \begin{pmatrix} \frac{4}{h^2}\sin^2\left(\frac{\xi h}{2}\right) & 0 \\ 0 & \frac{s-\cos^2\left(\frac{\xi h}{2}\right)}{h^2} \end{pmatrix}. \tag{4.31}$$

The system (4.15) can be transformed into the following Cauchy problem associated to a the system of two linear ODE's whose unknown is the vector function $\widehat{U}^h(\xi, t)$:

$$\begin{cases} M_h(\xi)\widehat{U}_{tt}^h(\xi, t) + R_h^s(\xi)\widehat{U}^h(\xi, t) = 0, & \xi \in \Pi_h,\ t > 0, \\ \widehat{U}^h(\xi, 0) = \widehat{U}^{h,0}(\xi), \quad \widehat{U}_t^h(\xi, 0) = \widehat{U}^{h,1}(\xi), & \xi \in \Pi_h. \end{cases} \tag{4.32}$$

Denote by $\langle\cdot,\cdot\rangle$ the canonical inner product in $\mathbb{C}^2$. Using the Parseval identity in (3.14), we

can write the total energy associated to (4.15) in its Fourier representation as follows:

$$\begin{aligned}E_h^s(\overrightarrow{U}^0,\overrightarrow{U}^1) &= \frac{1}{4\pi}\int\limits_{\Pi_h}\left[\langle M_h(\xi)\widehat{U}^{h,1}(\xi),\widehat{U}^{h,1}(\xi)\rangle + \langle R_h^s(\xi)\widehat{U}^{h,0}(\xi),\widehat{U}^{h,0}(\xi)\rangle\right]d\xi\\ &= \frac{1}{4\pi}\int\limits_{\Pi_h}\Big[\frac{1}{3}|\widehat{A}^{h,1}(\xi)|^2+\frac{2}{3}\Big|\frac{1}{2}\sin\big(\frac{\xi h}{2}\big)\widehat{J}^{h,1}(\xi)-i\cos\big(\frac{\xi h}{2}\big)\widehat{A}^{h,1}(\xi)\Big|^2+\frac{1}{12}|\widehat{J}^{h,1}(\xi)|^2\Big]d\xi\\ &+\frac{1}{4\pi}\int\limits_{\Pi_h}\Big[\frac{4}{h^2}\sin^2\big(\frac{\xi h}{2}\big)|\widehat{A}^{h,0}(\xi)|^2+\frac{s-\cos^2\left(\frac{\xi h}{2}\right)}{h^2}|\widehat{J}^{h,0}(\xi)|^2\Big]d\xi. \end{aligned}\tag{4.33}$$

The system of ODEs (4.32) can be written in the following simpler form

$$\widehat{U}_{tt}^h(\xi,t) = -S_h^s(\xi)\widehat{U}^h(\xi,t),\tag{4.34}$$

with

$$S_h^s(\xi) = M_h^{-1}(\xi)R_h^s(\xi) = \begin{pmatrix} \big(2-\cos(\xi h)\big)\frac{4}{h^2}\sin^2\left(\frac{\xi h}{2}\right) & -2i\sin(\xi h)\frac{s-\cos^2\left(\frac{\xi h}{2}\right)}{h^2}\\ 2i\sin(\xi h)\frac{4}{h^2}\sin^2\left(\frac{\xi h}{2}\right) & 4\big(2+\cos(\xi h)\big)\frac{s-\cos^2\left(\frac{\xi h}{2}\right)}{h^2}\end{pmatrix}.$$

Properties of the eigenvalues of the matrix $S_h^s(\xi)$. The eigenvalues Λ of the matrix $S_h^s(\xi)$ are the solutions of the following characteristic equation:

$$\Lambda^2 - 2\Lambda\frac{12+2(s-3)\big(2+\cos(\xi h)\big)}{h^2}+\frac{48}{h^4}\sin^2\big(\frac{\xi h}{2}\big)\Big(s-\cos^2\big(\frac{\xi h}{2}\big)\Big)=0.\tag{4.35}$$

With the notation

$$\Delta_h^s(\xi) = \big(12+2(s-3)(2+\cos(\xi h))\big)^2 - 48\sin^2\big(\frac{\xi h}{2}\big)\Big(s-\cos^2\big(\frac{\xi h}{2}\big)\Big),\tag{4.36}$$

the two roots in (4.35) are

$$\Lambda_{ph,h}^s(\xi) = \frac{12+2(s-3)(2+\cos(\xi h))-\sqrt{\Delta_h^s(\xi)}}{h^2},\quad \Lambda_{sp,h}^s(\xi) = \frac{12+2(s-3)(2+\cos(\xi h))+\sqrt{\Delta_h^s(\xi)}}{h^2}.\tag{4.37}$$

From now on, we will use several expressions involving $1/\Delta_h^s(\xi)$. The following proposition describes the pairs (ξ,s) for which $\Delta_h^s=0$.

Lemma 4.3.1. *The discriminant $\Delta_h^s(\xi)$ has the following property:*

$$\Delta_h^s(\xi)=0 \text{ iff } \Big(\xi=\pm\pi/h \text{ and } s=3\Big) \text{ or } \Big(\xi=0 \text{ and } s=1\Big).$$

Otherwise, $\Delta_h^s(\xi)>0$.

Proof. An easy computation shows that $\Delta_h^s(\xi)$ can be written as a sum of two squares:

$$\Delta_h^s(\xi) = \Big(2(s-3)(2+\cos(\xi h))+\frac{36\cos^2\left(\frac{\xi h}{2}\right)}{2+\cos(\xi h)}\Big)^2+\frac{192\cos^2\left(\frac{\xi h}{2}\right)\sin^6\left(\frac{\xi h}{2}\right)}{(2+\cos(\xi h))^2}.\tag{4.38}$$

The last term in the above identity is zero in the following two cases:

i. $\cos\left(\frac{\xi h}{2}\right)=0$. Under this condition, taking into account that $2+\cos(\xi h)>0$, the first term in (4.38) is zero iff $s=3$.

ii. $\sin\left(\frac{\xi h}{2}\right)=0$. Under this condition, the first term is zero iff $s=1$.

□

Following the terminology introduced in [11], [15], we call $\Lambda^s_{ph,h}(\xi)$, $\Lambda^s_{sp,h}(\xi)$ the physical (acoustical) and the spurious (optical) dispersion relation, respectively.

For a fixed $\xi \in \Pi_h$, $\Lambda^s_{ph,h}(\xi) \to \Lambda_h(\xi)$ as $s \to \infty$ and $\Lambda^s_{ph,h}(\xi) \to \Omega_h(\xi)$ as $s \to 1$, where $\Lambda_h(\xi)$ and $\Omega_h(\xi)$ are the dispersion relations corresponding to the P_1-classical finite element method and to the finite difference scheme, given by

$$\Lambda_h(\xi) = \frac{6}{h^2}\frac{1-\cos(\xi h)}{2+\cos(\xi h)} \text{ and } \Omega_h(\xi) = \frac{4}{h^2}\sin^2\Big(\frac{\xi h}{2}\Big). \tag{4.39}$$

The following result concerning the comparison between the dispersion relations $\Lambda^s_{ph,h}$, Λ_h and Ω_h holds:

Lemma 4.3.2. *For each $s \geqslant 1$ and $\xi \in \Pi_h$, the following inequalities hold:*

$$\Omega_h(\xi) \leqslant \Lambda^s_{ph,h}(\xi) \leqslant \Lambda_h(\xi). \tag{4.40}$$

Proof. Firstly, we prove the second inequality in (4.40). It is easy to check that

$$\begin{aligned}\Delta^s_h(\xi) - \big|4(2+\cos(\xi h))\big(s-\cos^2\big(\frac{\xi h}{2}\big)\big) - \big(12+2(s-3)(2+\cos(\xi h))\big)\big|^2 \\ = 64\sin^4\Big(\frac{\xi h}{2}\Big)\cos^2\Big(\frac{\xi h}{2}\Big)\Big(s-\cos^2\Big(\frac{\xi h}{2}\Big)\Big) \geq 0.\end{aligned}$$

The last inequality implies that

$$\frac{\Lambda^s_{ph,h}(\xi)}{3\Omega_h(\xi)} = \frac{4\big(s-\cos^2\big(\frac{\xi h}{2}\big)\big)}{12+2(s-3)(2+\cos(\xi h))+\sqrt{\Delta^s_h(\xi)}} \leqslant \frac{1}{2+\cos(\xi h)} = \frac{\Lambda_h(\xi)}{3\Omega_h(\xi)}.$$

Using the second inequality in (4.40) proved before, we get

$$\partial_s\Lambda^s_{ph,h}(\xi) = \frac{2(2+\cos(\xi h))}{\sqrt{\Delta^s_h(\xi)}}\big[\Lambda_h(\xi) - \Lambda^s_{ph,h}(\xi)\big] \geqslant 0, \tag{4.41}$$

i.e., $\Lambda^s_{ph,h}(\xi)$ is increasing with respect to s. The maximum value is attained as s goes to infinity and is equal to $\Lambda_h(\xi)$ and the minimum value is attained for $s=1$ and is equal to $\Omega_h(\xi)$. This concludes the proof of (4.40). □

At the same time, $\Lambda^s_{sp,h}(\xi) \to \infty$ as $s \to \infty$.

For all $\xi \in \Pi_h$, set

$$\lambda^s_{ph,h}(\xi) = \text{sign}(\xi)\sqrt{\Lambda^s_{ph,h}(\xi)}, \quad \lambda^s_{sp,h}(\xi) = \text{sign}(\xi)\sqrt{\Lambda^s_{sp,h}(\xi)},$$

and

$$\lambda_h(\xi) = \text{sign}(\xi)\sqrt{\Lambda_h(\xi)}, \quad \omega_h(\xi) = \text{sign}(\xi)\sqrt{\Omega_h(\xi)}.$$

In Figure 4.5, we represent the physical dispersion relation $\lambda^s_{ph,1}(\xi)$ (black) and the spurious one $\lambda^s_{sp,1}(\xi)$ (dotted black) for $s=1.5$ (top, left), $s=2$ (top, right), $s=3$ (bottom, left), $s=5$ (bottom, right) compared to the dispersion relation corresponding to the continuous wave equation ξ (blue) and to those corresponding to its finite-difference $\omega_1(\xi)$ (red) and classical P_1-finite element $\lambda_1(\xi)$ semi-discretizations (green). The marked points correspond to wave numbers where the corresponding group velocity vanishes.

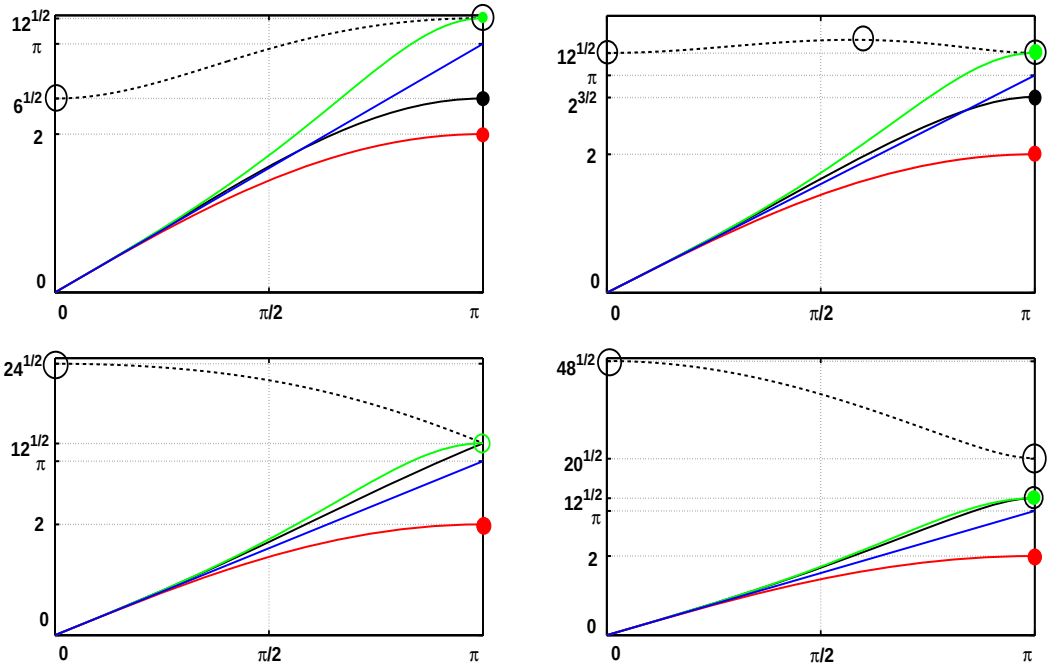

Figure 4.5: Physical dispersion relation (black) versus the spurious one (dotted black) for the SIPG semi-discrete wave equation.

Properties of the eigenvectors of the matrix $S_h^s(\xi)$ in (4.32). The two eigenvectors of $S_h^s(\xi)$ are:

$$P_{ph,h}^s(\xi) = \frac{1}{\sqrt{1+|p_{ph,h}^s(\xi)|^2}} \begin{pmatrix} 1 \\ p_{ph,h}^s(\xi) \end{pmatrix}, \quad P_{sp,h}^s(\xi) = \frac{1}{\sqrt{1+|p_{sp,h}^s(\xi)|^2}} \begin{pmatrix} p_{sp,h}^s(\xi) \\ 1 \end{pmatrix}, \quad (4.42)$$

where

$$p_{ph,h}^s(\xi) = \frac{1}{i} \frac{\sin\left(\frac{\xi h}{2}\right)(2-\cos(\xi h)) - \frac{12\sin\left(\frac{\xi h}{2}\right)\left(s-\cos^2\left(\frac{\xi h}{2}\right)\right)}{12+2(s-3)(2+\cos(\xi h))+\sqrt{\Delta_h^s(\xi)}}}{\cos\left(\frac{\xi h}{2}\right)\left(s-\cos^2\left(\frac{\xi h}{2}\right)\right)} \quad (4.43)$$

and

$$p_{sp,h}^s(\xi) = \frac{1}{i} \frac{2-\cos(\xi h) - \frac{12\left(s-\cos^2\left(\frac{\xi h}{2}\right)\right)}{12+2(s-3)(2+\cos(\xi h))+\sqrt{\Delta_h^s(\xi)}}}{2\sin(\xi h)}. \quad (4.44)$$

The following two propositions describe the behavior of the two eigenvectors $P_{ph,h}^s(\xi)$ and $P_{sp,h}^s(\xi)$ of $S_h^s(\xi)$.

Proposition 4.3.2. *The physical eigenvector $P_{ph,h}^s(\xi)$ has the following properties:*

Ph1. For all $s>1$, $\lim_{\xi\to 0} p_{ph,h}^s(\xi) = 0$ and $\lim_{\xi\to 0} P_{ph,h}^s(\xi) = \begin{pmatrix} 1 & 0 \end{pmatrix}^T$.

Ph2. For all $s>3$, $\lim_{\xi\to \pm\pi/h} p_{ph,h}^s(\xi) = 0$ and $\lim_{\xi\to \pm\pi/h} P_{ph,h}^s(\xi) = \begin{pmatrix} 1 & 0 \end{pmatrix}^T$.

Ph3. $\lim_{\xi\to \pm\pi/h} p_{ph,h}^3(\xi) = \pm\frac{2}{i\sqrt{3}}$ and $\lim_{\xi\to \pm\pi/h} P_{ph,h}^3(\xi) = \begin{pmatrix} \frac{\sqrt{3}}{\sqrt{7}} & \pm\frac{2}{i\sqrt{7}} \end{pmatrix}^T$.

Ph4. For all $s \in (1,3)$, $\lim_{\xi \to \pm\pi/h} ip^s_{ph,h}(\xi) = \pm\infty$ *and* $\lim_{\xi \to \pm\pi/h} P^s_{ph,h}(\xi) = \begin{pmatrix} 0 & \pm\frac{1}{i} \end{pmatrix}^T$.

In Figure 4.6, we represent a discrete function $(F_j)_{j\in\mathbb{Z}} = (A_j, J_j)^T_{j\in\mathbb{Z}}$ whose Fourier representation is $\widehat{F}^h(\xi) = (\widehat{A}^h(\xi), \widehat{J}^h(\xi))^T = 2\pi P^s_{ph,h}(\xi)\delta_{\xi_0}(\xi)$. We represent $(\mathrm{Re}(F), \mathrm{Im}(F))$ (blue, respectively red) for the pairs $(s > 1, \xi_0 = 0)$ (top, left), $(s > 3, \xi_0 = \pi/h)$ (top, right), $(s = 3, \xi_0 = \pi/h)$ (bottom, left) and $(s \in (1,3), \xi_0 = \pi/h)$ (bottom, right).

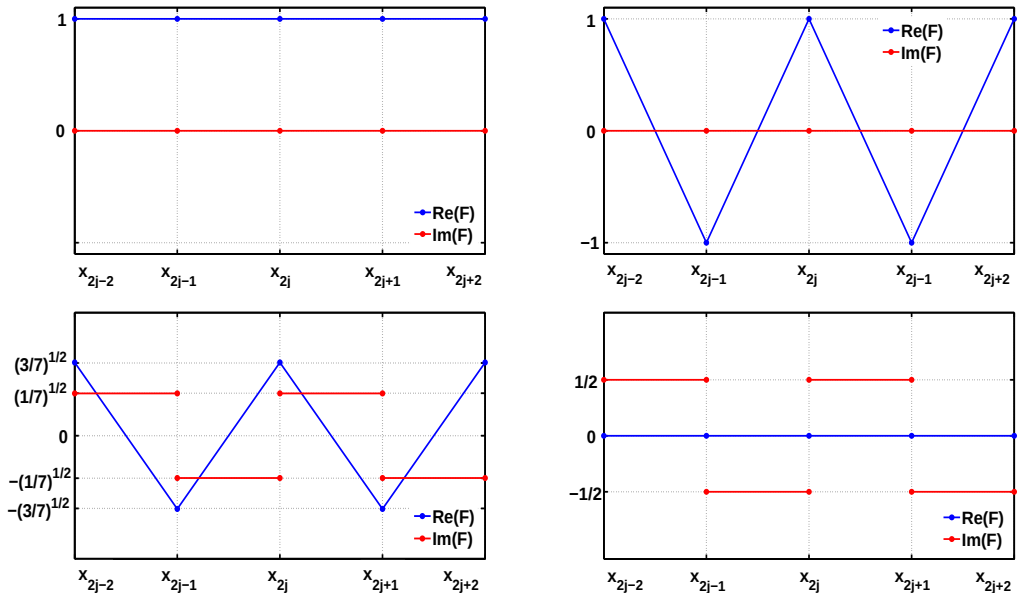

Figure 4.6: Real part versus imaginary part of the physical eigenvector at the critical wave numbers.

Proof. From the expression (4.43), observe that $p^s_{ph,h}(\xi)$ could be singular only for $\xi = \pm\pi/h$. Then the property (Ph1) concerning the behavior of the eigenvector $P^s_{ph,h}(\xi)$ in $\xi = 0$ is just an easy computation. In order to see if $p^s_{ph,h}(\xi)$ is singular in $\xi = \pm\pi/h$, we firstly study the limit as $\xi \to \pm\pi/h$ of the numerator in (4.43). In fact,

$$\lim_{\xi\to\pm\pi/h}\Big[\sin\big(\frac{\xi h}{2}\big)(2-\cos(\xi h)) - \frac{12\sin\big(\frac{\xi h}{2}\big)\big(s-\cos^2\big(\frac{\xi h}{2}\big)\big)}{12+2(s-3)(2+\cos(\xi h))+\sqrt{\Delta^s_h(\xi)}}\Big] = \begin{cases} \pm(3-s), & s<3 \\ 0, & s\geq 3. \end{cases}$$

The property (Ph4) follows directly from the above limit. In order to obtain (Ph2), one applies L'Hôpital Rule.

According to Lemma 4.3.1, the behavior of $p^3_{ph,h}(\xi)$ as $\xi \to \pm\pi/h$ has to be considered separately. The property (Ph3) follows from the fact that $p^3_{ph,h}(\xi)$ takes the form

$$ip^3_{ph,h}(\xi) = \frac{\sin\big(\frac{\xi h}{2}\big)\big[-3\cos\big(\frac{\xi h}{2}\big)+(2-\cos(\xi h))\sqrt{9+3\sin^2\big(\frac{\xi h}{2}\big)}\big]}{\big(2+\sin^2\big(\frac{\xi h}{2}\big)\big)\big(3+\cos\big(\frac{\xi h}{2}\big)\sqrt{9+3\sin^2\big(\frac{\xi h}{2}\big)}\big)}.$$

□

Proposition 4.3.3. *The spurious eigenvector* $P^s_{sp,h}(\xi)$ *has the following properties:*

Sp1. For all $s > 1$, $\lim_{\xi\to 0} p^s_{sp,h}(\xi) = 0$ *and* $\lim_{\xi\to 0} P^s_{sp,h}(\xi) = \begin{pmatrix} 0 & 1 \end{pmatrix}^T$.

Sp2. For all $s > 3$, $\lim_{\xi\to \pm\pi/h} ip^s_{sp,h}(\xi) = 0$ *and* $\lim_{\xi\to \pm\pi/h} P^s_{sp,h}(\xi) = \begin{pmatrix} 0 & 1 \end{pmatrix}^T$.

Sp3. $\lim_{\xi\to \pm\pi/h} ip^3_{sp,h}(\xi) = \pm\sqrt{3}/2$ *and* $\lim_{\xi\to \pm\pi/h} P^3_{sp,h}(\xi) = \begin{pmatrix} \pm\frac{\sqrt{3}}{i\sqrt{7}} & \frac{2}{\sqrt{7}} \end{pmatrix}^T$.

Sp4. For all $s \in (1,3)$, $\lim_{\xi\to \pm\pi/h} ip^s_{sp,h}(\xi) = \pm\infty$ *and* $\lim_{\xi\to \pm\pi/h} P^s_{sp,h}(\xi) = \begin{pmatrix} \pm\frac{1}{i} & 0 \end{pmatrix}^T$.

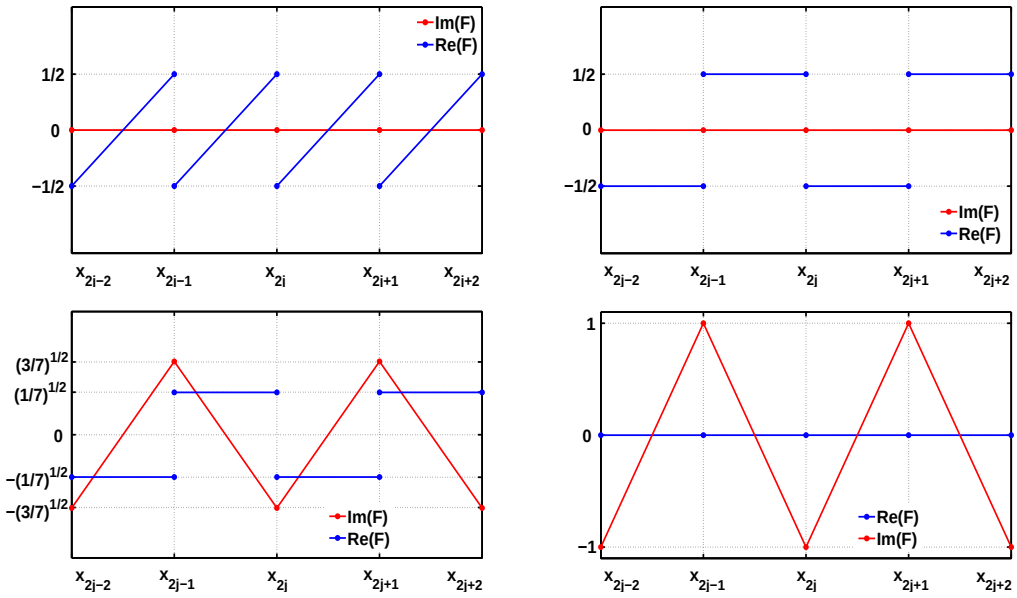

Figure 4.7: Real part versus imaginary part of the spurious eigenvector at the critical wave numbers.

In Figure 4.7, we represent a discrete function $(F_j)_{j\in\mathbb{Z}} = (A_j, J_j)^T_{j\in\mathbb{Z}}$ whose Fourier representation is $\widehat{F}^h(\xi) = (\widehat{A}^h(\xi), \widehat{J}^h(\xi))^T = 2\pi P^s_{sp,h}(\xi)\delta_{\xi_0}(\xi)$. We represent $(\mathrm{Re}(F), \mathrm{Im}(F))$ (blue, respectively red) for the pairs $(s > 1, \xi_0 = 0)$ (top, left), $(s > 3, \xi_0 = \pi/h)$ (top, right), $(s = 3, \xi_0 = \pi/h)$ (bottom, left) and $(s \in (1,3), \xi_0 = \pi/h)$ (bottom, right).

Proof. From (4.44), we observe that $p^s_{sp,h}(\xi)$ could be singular for $\xi = 0$ and $\xi = \pm\pi/h$.

The case $\xi = 0$**.** Observe that

$$\lim_{\xi\to 0}\left[2 - \cos(\xi h) - \frac{12\big(s - \cos^2\left(\frac{\xi h}{2}\right)\big)}{12 + 2(s-3)(2+\cos(\xi h)) + \sqrt{\Delta^s_h(\xi)}}\right] = 0.$$

The property (Sp1) follows from L'Hôpital Rule.

The case $\xi = \pm\pi/h$**.** In order to see if $p^s_{sp,h}(\xi)$ is singular as $\xi \to \pm\pi/h$, we study the limit as $\xi \to \pm\pi/h$ of the numerator in (4.44). In fact,

$$\lim_{\xi\to \pm\pi/h}\left[2 - \cos(\xi h) - \frac{12\big(s - \cos^2\left(\frac{\xi h}{2}\right)\big)}{12 + 2(s-3)(2+\cos(\xi h)) + \sqrt{\Delta^s_h(\xi)}}\right] = \begin{cases} 3 - s, & s < 3 \\ 0, & s \geq 3. \end{cases}$$

Then the property (Sp4) concerning the behavior of $P^s_{sp,h}(\xi)$ as $\xi \to \pm\pi/h$ and $s < 3$ follows directly from the above limit. In order to obtain (Sp2), one has to apply L'Hôpital Rule.

According to Lemma 4.3.1, the case $s = 3$ and $\xi \to \pm\pi/h$ has to be considered separately. The property (Sp3) follows from the fact when $s = 3$, $p^3_{sp,h}(\xi)$ has the following form:

$$ip^3_{sp,h}(\xi) = \frac{3\sin\left(\frac{\xi h}{2}\right)\left(2+\sin^2\left(\frac{\xi h}{2}\right)\right)^2}{\left[3+\cos\left(\frac{\xi h}{2}\right)\sqrt{9+3\sin^2\left(\frac{\xi h}{2}\right)}\right]\left[3\cos\left(\frac{\xi h}{2}\right)+\left(1+2\sin^2\left(\frac{\xi h}{2}\right)\right)\sqrt{9+3\sin^2\left(\frac{\xi h}{2}\right)}\right]}.$$

□

We denote by $P^s_h(\xi)$ the 2×2-matrix whose columns are the eigenvectors $P^s_{ph,h}(\xi)$ and $P^s_{sp,h}(\xi)$ and by $(P^s_h(\xi))^{-1}$ its inverse. Namely,

$$P^s_h(\xi) = \begin{pmatrix} \frac{1}{\sqrt{1+|p^s_{ph,h}(\xi)|^2}} & \frac{p^s_{sp,h}(\xi)}{\sqrt{1+|p^s_{sp,h}(\xi)|^2}} \\ \frac{p^s_{ph,h}(\xi)}{\sqrt{1+|p^s_{ph,h}(\xi)|^2}} & \frac{1}{\sqrt{1+|p^s_{sp,h}(\xi)|^2}} \end{pmatrix}. \tag{4.45}$$

The matrix $S^s_h(\xi)$ can be decomposed as follows

$$S^s_h(\xi) = P^s_h(\xi)\begin{pmatrix} \Lambda^s_{ph,h}(\xi) & 0 \\ 0 & \Lambda^s_{sp,h}(\xi) \end{pmatrix}\left(P^s_h(\xi)\right)^{-1}. \tag{4.46}$$

The solution of the problem (4.32) is given by

$$\begin{aligned}\widehat{U}^h(\xi,t) = \sum_{\pm}\frac{1}{2}\Big[&P^s_h(\xi)\begin{pmatrix} \exp(\pm it\lambda^s_{ph,h}(\xi)) & 0 \\ 0 & \exp(\pm it\lambda^s_{sp,h}(\xi)) \end{pmatrix}\left(P^s_h(\xi)\right)^{-1}\widehat{U}^{h,0}(\xi) \\ &+ P^s_h(\xi)\begin{pmatrix} \pm\frac{\exp(\pm it\lambda^s_{ph,h}(\xi))}{i\lambda^s_{ph,h}(\xi)} & 0 \\ 0 & \pm\frac{\exp(\pm it\lambda^s_{sp,h}(\xi))}{i\lambda^s_{sp,h}(\xi)} \end{pmatrix}\left(P^s_h(\xi)\right)^{-1}\widehat{U}^{h,1}(\xi)\Big].\end{aligned} \tag{4.47}$$

The analysis of the group velocities. The group velocity corresponding to the physical dispersion relation $\lambda^s_{ph,h}(\xi)$ takes the following form:

$$\partial_\xi\lambda^s_{ph,h}(\xi) = \sqrt{\frac{12+2(s-3)(2+\cos(\xi h))+\sqrt{\Delta^s_h(\xi)}}{12\left(s-\cos^2\left(\frac{\xi h}{2}\right)\right)}}\frac{\cos\left(\frac{\xi h}{2}\right)}{\sqrt{\Delta^s_h(\xi)}}e^s_{ph,h}(\xi),$$

with

$$e^s_{ph,h}(\xi) = (s-3)\Big(12+2(s-3)(2+\cos(\xi h))-\sqrt{\Delta^s_h(\xi)}\Big)+6(s-\cos(\xi h)). \tag{4.48}$$

When $s = 3$, the group velocity corresponding to $\lambda^3_{ph,h}(\xi)$ takes the form

$$\partial_\xi\lambda^3_{ph,h}(\xi) = \frac{\left(1+\sin^2\left(\frac{\xi h}{2}\right)\right)\sqrt{3+\cos\left(\frac{\xi h}{2}\right)\sqrt{9+3\sin^2\left(\frac{\xi h}{2}\right)}}}{\sqrt{2+\sin^2\left(\frac{\xi h}{2}\right)}\sqrt{3+\sin^2\left(\frac{\xi h}{2}\right)}}.$$

Proposition 4.3.4. *The group velocity $\partial_\xi\lambda^s_{ph,h}(\xi)$ has the following properties:*

GVPh1. For all $s > 1$, $\lim_{\xi\to 0}\partial_\xi\lambda^s_{ph,h}(\xi) = 1$.

GVPh2. For all $s \in (1,\infty)\setminus\{3\}$, $\lim_{\xi\to\pm\pi/h}\partial_\xi\lambda^s_{ph,h}(\xi) = 0$.

GVPh3. $\lim_{\xi\to\pm\pi/h} \partial_\xi \lambda^3_{ph,h}(\xi) = 1.$

GVPh4. For all $s \in (1,\infty)$ *and all* $\xi \in \Pi_h$, $e^s_{ph,h}(\xi) > 0.$

Proof. The properties (GVPh1-GVPh3) follow directly from the explicit expression of the group velocity $\partial_\xi \lambda^s_{ph,h}(\xi)$. The property (GVPh1) is related to the convergence of the SIPG approximation for the wave equation as $h \to 0$ to the continuous wave equation, for which the group velocity is identically one. The property (GVPh2) is usual for the group velocity corresponding to the classical finite difference and P_1-finite element approximations of the wave equation. Its consequence is that it is possible to construct wave packets concentrated on the physical dispersion diagram propagating at arbitrarily low speed. The property (GVPh3) is in some sense unusual, but it is related to the singular behavior of $\Delta^3_h(\xi)$ as $\xi \to \pm\pi/h$ proved in Lemma 4.3.1. The following identity proves the property (GVPh4):

$$e^s_{ph,h}(\xi) = \frac{12(s-1)\big[(s-1)\big(3+2\sin^2\left(\frac{\xi h}{2}\right)\big)+2\sin^2\left(\frac{\xi h}{2}\right)\big]+6(s-\cos(\xi h))\sqrt{\Delta^s_h(\xi)}}{12+2(s-3)(2+\cos(\xi h))+\sqrt{\Delta^s_h(\xi)}} > 0.$$

The property (GVPh4) means that the unique singular point of $\lambda^s_{ph,h}(\xi)$ is $\xi = \pm\pi/h$ and that $\partial_\xi \lambda^s_{ph,h}(\xi) \geq 0$, i.e. $\lambda^s_{ph,h}(\xi)$ is strictly increasing in ξ. □

Remark 4.3.1. *A consequence of the fact that* $\lambda^s_{ph,h}(\xi)$ *is increasing in* ξ *is the following one:*

$$\max_{\xi\in\Pi_h} |\lambda^s_{ph,h}(\xi)| = |\lambda^s_{ph,h}(\pm\pi/h)| = \begin{cases} \frac{2\sqrt{s}}{h}, & s \in (1,3) \\ \frac{2\sqrt{3}}{h}, & s \geq 3. \end{cases}$$

The group velocity corresponding to the spurious dispersion relation $\lambda^s_{sp,h}(\xi)$ takes the following form:

$$\partial_\xi \lambda^s_{sp,h}(\xi) = -\frac{\sin(\xi h)}{\sqrt{12+2(s-3)(2+\cos(\xi h))+\sqrt{\Delta^s_h(\xi)}}\sqrt{\Delta^s_h(\xi)}} e^s_{sp,h}(\xi),$$

with

$$e^s_{sp,h}(\xi) = (s-3)\big(12+2(s-3)(2+\cos(\xi h))+\sqrt{\Delta^s_h(\xi)}\big)+6(s-\cos(\xi h)). \tag{4.49}$$

When $s = 3$, the group velocity corresponding to $\lambda^3_{sp,h}(\xi)$ takes the form

$$\partial_\xi \lambda^3_{ph,h}(\xi) = -\frac{3\sin\left(\frac{\xi h}{2}\right)\left(1+\sin^2\left(\frac{\xi h}{2}\right)\right)}{\sqrt{3+\cos\left(\frac{\xi h}{2}\right)\sqrt{9+3\sin^2\left(\frac{\xi h}{2}\right)}}\sqrt{9+3\sin^2\left(\frac{\xi h}{2}\right)}}.$$

Proposition 4.3.5. *The group velocity* $\partial_\xi \lambda^s_{sp,h}(\xi)$ *has the following properties:*

GVSp1 For all $s > 1$, $\lim_{\xi\to 0} \partial_\xi \lambda^s_{sp,h}(\xi) = 0.$

GVSp2. For all $s \in (1,\infty)\setminus\{3\}$, $\lim_{\xi\to\pm\pi/h} \partial_\xi \lambda^s_{sp,h}(\xi) = 0.$

GVSp3. $\lim_{\xi\to\pm\pi/h} \partial_\xi \lambda^3_{sp,h}(\xi) = -1.$

GVSp4. For all $s \in (5/2,\infty)$ *and all* $\xi \in \Pi_h$, $e^s_{sp,h}(\xi) > 0.$

GVSp5. For all $s \in (1,5/3)$ *and all* $\xi \in \Pi_h$, $e^s_{sp,h}(\xi) < 0.$

GVSp6. For all $s \in (5/3,5/2)$ *and all* $\xi \in \Pi_h$, $e^s_{sp,h}(0) < 0$ *and* $e^s_{sp,h}(\pi/h) > 0.$

Remark 4.3.2. *The spourious diagram $\lambda^s_{sp,h}(\xi)$ provides a rich behaviour from the point of view of monotonicity, according to several ranges of the penalization parameter s, as follows:*

- For $s \geq 3$, $\lambda^s_{sp,h}(\xi)$ is decreasing on $(0,\pi/h)$, with minimal value in $\xi = \pi/h$ equal to $4\sqrt{s}/h$ and the maximal one in $\xi = 0$, equal to $2\sqrt{3(s-1)}/h$.

- For $s \in (5/2,3)$, $\lambda^s_{sp,h}(\xi)$ is decreasing on $(0,\pi/h)$, with minimal value in $\xi = \pi/h$ equal to $2\sqrt{3}/h$ and the maximal one in $\xi = 0$, equal to $2\sqrt{3(s-1)}/h$.

- For $s \in (2,5/2)$, $\lambda^s_{sp,h}(\xi)$ has a maximum point on $(0,\pi/h)$, with minimal value in $\xi = \pi/h$ equal to $2\sqrt{3}/h$ and the maximal one in $\xi = 0$, equal to $2\sqrt{3(s-1)}/h$.

- For $s \in (5/3,2)$, $\lambda^s_{sp,h}(\xi)$ has a maximum point on $(0,\pi/h)$, with maximal value in $\xi = \pi/h$ equal to $2\sqrt{3}/h$ and the minimal one in $\xi = 0$, equal to $2\sqrt{3(s-1)}/h$.

- For $s \in (1,5/3)$, $\lambda^s_{sp,h}(\xi)$ is strictly increasing $(0,\pi/h)$, with maximal value in $\xi = \pi/h$ equal to $2\sqrt{3}/h$ and the minimal one in $\xi = 0$, equal to $2\sqrt{3(s-1)}/h$.

Concerning the behavior of $e^s_{sp,h}(\xi)$ for $s \in (5/3,5/2)$, the following more precise result holds:

Proposition 4.3.6. *For $s \in (5/3,5/2)$, there exists an unique solution $\xi_s h \in (0,\pi)$ of the equation $e^s_{sp,h}(\xi) = 0$, explicitly given by*

$$\cos(\xi_s h) = \frac{s(2s-3) - (s-1)(3-s)\sqrt{6(s-1)}}{3-(3-s)^2}. \tag{4.50}$$

Proof. Taking into account the fact that $\Lambda^s_{sp,h}(\xi) \geq 0$ for all $s \geq 1$ and $\xi \in \Pi_h$, we see that for $s \geq 3$, $e^s_{sp,h}(\xi) = 0$ has no solutions. This follows from the fact that $e^s_{sp,h}(\xi)$ can be written as

$$e^s_{sp,h}(\xi) = (s-3)h^2\Lambda^s_{sp,h}(\xi) + 6(s - \cos(\xi h)).$$

Finding solutions of $e^s_{sp,h}(\xi) = 0$ is equivalent to finding solutions of

$$f^s_{sp,h}(\xi) = (3-s)\sqrt{\Delta^s_h(\xi)}, \tag{4.51}$$

with

$$f^s_{sp,h}(\xi) = 6(s - \cos(\xi h)) - (3-s)\big(12 + 2(s-3)(2+\cos(\xi h))\big).$$

Taking into account that, for $s \in (1,3)$, the right hand side in the equation (4.51) is positive, in order to guarantee the existence of a solution of (4.51), we have to find the values of $(\xi h, s) \in (0,\pi)\times(1,3)$ such that $f^s_{sp,h}(\xi) \geq 0$.

We eliminate some values of $s \in (1,3)$ as a consequence of the monotonicity of $f^s_{sp,h}(\xi)$ in ξ. To study the monotonicity, remark that the first order derivative of $f^s_{sp,h}(\xi)$ is given by

$$\partial_\xi f^s_{sp,h}(\xi) = 2h\sin(\xi h)\big(3-(s-3)^2\big).$$

We analyze two cases:

- For all $s \in (1, 3-\sqrt{3})$, $f^s_{sp,h}(\xi)$ is decreasing, so that $0 > f^s_{sp,h}(0) = 6(s-1)(s-2) \geq f^s_{sp,h}(\xi) \geq f^s_{sp,h}(\pi/h) = 2s^2+6s-12$, for all $\xi h \in (0,\pi)$. Then, for $s \in (1,3-\sqrt{3})$, there are not solutions of the equation $f^s_{sp,h}(\xi) = 0$

- For all $s \in (3-\sqrt{3},3)$, $f^s_{sp,h}(\xi)$ is increasing, so that $f^s_{sp,h}(0) = 6(s-1)(s-2) \leq f^s_{sp,h}(\xi) \leq f^s_{sp,h}(\pi/h) = 2s^2+6s-12$. For $s \in (3-\sqrt{3},(\sqrt{33}-3)/2)$, $f^s_{sp,h}(\xi) \leq f^s_{sp,h}(\pi/h) < 0$, for all $\xi h \in (0,\pi)$, so that $e^s_{sp,h}(\xi) = 0$ has no solution. In the following, we focus on $s \in ((\sqrt{33}-3)/2,3)$.

By replacing $\cos(\xi h) = X$ in (4.51), we see that X is a solution of the quadratic equation

$$(3-(3-s)^2)X^2 - 2s(2s-3)X + 6s^3 - 22s^2 + 24s - 9 = 0, \tag{4.52}$$

whose solutions are explicitly given by

$$X^{\pm}(s) = \frac{s(2s-3) \pm (s-1)(3-s)\sqrt{6(s-1)}}{3-(3-s)^2}.$$

The values of $s \in ((\sqrt{33}-3)/2, 3)$ for which there exist solutions of (4.51) coincide with those for which at least one of the two numbers $X^{\pm}(s)$ belongs to $[-1,1]$. When this happens, the solution $\xi_s h \in (0,\pi)$ of (4.50) can founded as $\xi_s h = \arccos(X^{\pm}(s))$ and satisfying the restriction $f^s_{sp,h}(\xi_s) \geq 0$. The requirement $-1 \leq X^{\pm}(s) \leq 1$ is equivalent to

$$a(s) = -s^2 - 3s + 6 \leq \pm\delta(s) \leq b(s) = -3(s-1)(s-2), \text{ with } \delta(s) = (s-1)(3-s)\sqrt{6(s-1)}.$$

For $s \in (2,3)$, $a(s), b(s) < 0$, so that from the two values $\pm\delta(s)$, we choose only the negative one. Both inequalities $-a(s) \geq \delta(s) \geq -b(s)$ are verified iff $s \in (2, 5/2)$.

For $s \in ((\sqrt{33}-3)/2, 2)$, $a(s) < 0$ and $b(s) > 0$, so that we verify under which conditions on s one of the inequalities $a(s) \leq -\delta(s)$ or $\delta(s) \leq b(s)$ holds. The second one implies $s \geq 5/2$, so it has no solutions $s \in ((\sqrt{33}-3)/2, 2)$. The first one implies $s \geq 5/3$, so that $s \in (5/3, 2)$. Consequently, for all $s \in (5/3, 5/2)$, $X^-(s) \in (-1,1)$ and there exists an unique solution $\xi_s h \in (0,\pi)$ such that $\cos(\xi_s h) = X^-(s)$. It is easy to verify that $f^s_{sp,h}(\xi_s) = 2(s-1)(3-s)\sqrt{6(s-1)} > 0$. □

4.4 Concentrated waves for the SIPG method

Choice of the Fourier mode. In this section we will focus on initial data with one of the following two properties:

$$\widehat{U}^{h,i}(\xi) = P^s_{ph,h}(\xi)\widehat{u}^{h,i}(\xi) \tag{4.53}$$

or

$$\widehat{U}^{h,i}(\xi) = P^s_{sp,h}(\xi)\widehat{u}^{h,i}(\xi), \tag{4.54}$$

for $i = 0, 1$, where $\widehat{u}^{h,i}(\xi)$ are scalar functions defined on Π_h.

Solution and energy for initial data concentrated on the physical mode. For initial data in (4.32) with the property (4.53), the solution becomes concentrated only on the physical mode. Indeed, since $(P^s_h(\xi))^{-1}P^s_{ph,h}(\xi) = (1,0)^T$, the solution of (4.32), given initially by (4.47), simplifies to

$$\widehat{U}^h(\xi,t) = P^s_{ph,h}(\xi)\frac{1}{2}\sum_{\pm}\Big(\widehat{u}^{h,0}(\xi) \pm \frac{\widehat{u}^{h,1}(\xi)}{i\lambda^s_{ph,h}(\xi)}\Big)\exp(\pm it\lambda^s_{ph,h}(\xi)). \tag{4.55}$$

The associated total energy also simplifies with respect to the one indicated by (4.33):

$$E^s_h(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}) = \frac{1}{4\pi}\int_{\Pi_h}\left(m^s_{ph,h}(\xi)|\widehat{u}^{h,1}(\xi)|^2 + r^s_{ph,h}(\xi)|\widehat{u}^{h,0}(\xi)|^2\right)d\xi, \tag{4.56}$$

where

$$\begin{aligned} m^s_{ph,h}(\xi) &= \langle M_h(\xi)P^s_{ph,h}(\xi), P^s_{ph,h}(\xi)\rangle = \\ &= \frac{1}{1+|p^s_{ph,h}(\xi)|^2}\Big[\frac{1}{3} + \frac{1}{12}|p^s_{ph,h}(\xi)|^2 + \frac{2}{3}\Big|\cos\big(\frac{\xi h}{2}\big) + \frac{i}{2}\sin\big(\frac{\xi h}{2}\big)p^s_{ph,h}(\xi)\Big|^2\Big] \end{aligned}$$

and

$$r^s_{ph,h}(\xi) = \langle R^s_h(\xi)P^s_{ph,h}(\xi), P^s_{ph,h}(\xi)\rangle$$
$$= \frac{1}{1+|p^s_{ph,h}(\xi)|^2}\Big[\frac{4}{h^2}\sin^2\big(\frac{\xi h}{2}\big) + |p^s_{ph,h}(\xi)|^2\frac{s-\cos^2\left(\frac{\xi h}{2}\right)}{h^2}\Big].$$

Solution and energy for initial data concentrated on the spurious mode. For initial data in (4.32) with the property (4.54), the solution becomes concentrated only on the spurious mode. Indeed, since $(P^s_h(\xi))^{-1}P^s_{sp,h}(\xi) = (0,1)^T$, the solution of (4.32), given initially by (4.47), simplifies to

$$\widehat{U}^h(\xi,t) = P^s_{sp,h}(\xi)\frac{1}{2}\sum_{\pm}\Big(\widehat{u}^{h,0}(\xi) \pm \frac{\widehat{u}^{h,1}(\xi)}{i\lambda^s_{sp,h}(\xi)}\Big)\exp(\pm it\lambda^s_{sp,h}(\xi)). \tag{4.57}$$

The associated total energy also simplifies with respect to (4.33):

$$E^s_h(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}) = \frac{1}{4\pi}\int\limits_{\Pi_h}\big(m^s_{sp,h}(\xi)|\widehat{u}^{h,1}(\xi)|^2 + r^s_{sp,h}(\xi)|\widehat{u}^{h,0}(\xi)|^2\big)\,d\xi, \tag{4.58}$$

where

$$m^s_{sp,h}(\xi) = \langle M_h(\xi)P^s_{sp,h}(\xi), P^s_{sp,h}(\xi)\rangle =$$
$$= \frac{1}{1+|p^s_{sp,h}(\xi)|^2}\Big[\frac{1}{3}|p^s_{sp,h}(\xi)|^2 + \frac{1}{12} + \frac{2}{3}\Big|\cos\big(\frac{\xi h}{2}\big)p^s_{sp,h}(\xi) + \frac{i}{2}\sin\big(\frac{\xi h}{2}\big)\Big|^2\Big]$$

and

$$r^s_{sp,h}(\xi) = \langle R^s_h(\xi)P^s_{sp,h}(\xi), P^s_{sp,h}(\xi)\rangle$$
$$= \frac{1}{1+|p^s_{sp,h}(\xi)|^2}\Big[\frac{4}{h^2}\sin^2\big(\frac{\xi h}{2}\big)|p^s_{sp,h}(\xi)|^2 + \frac{s-\cos^2\left(\frac{\xi h}{2}\right)}{h^2}\Big].$$

Lemma 4.4.1. *For all $\xi \in \Pi_h$ and all $s \in (1,\infty)$, the following two identities hold:*

$$r^s_{ph,h}(\xi) = \Lambda^s_{ph,h}(\xi)m^s_{ph,h}(\xi) \text{ and } r^s_{sp,h}(\xi) = \Lambda^s_{sp,h}(\xi)m^s_{sp,h}(\xi). \tag{4.59}$$

Proof. We prove only the first identity, for the second one the arguments are the same.

Since $P^s_{ph,h}(\xi)$ is the eigenvector of the matrix $A^s_h(\xi) = (M_h(\xi))^{-1}R^s_h(\xi)$ corresponding to the eigenvalue $\Lambda^s_{ph,h}(\xi)$, we have

$$r^s_{ph,h}(\xi) = \langle M_h(\xi)A^s_h(\xi)P^s_{ph,h}(\xi), P^s_{ph,h}(\xi)\rangle = \Lambda^s_{ph,h}(\xi)\langle M_h(\xi)P^s_{ph,h}(\xi), P^s_{ph,h}(\xi)\rangle$$
$$= \Lambda^s_{ph,h}(\xi)m^s_{ph,h}(\xi).$$

□

Choice of the direction. For initial data with the property (4.53), we consider another restriction:

$$\widehat{u}^{h,1}(\xi) = i\lambda^s_{ph,h}(\xi)\widehat{u}^{h,0}(\xi).$$

With this second restriction, the solution simplifies with respect to the one given by (4.55):

$$\widehat{U}^h(\xi,t) = P^s_{ph,h}(\xi)\widehat{u}^{h,0}(\xi)\exp(it\lambda^s_{ph,h}(\xi)). \tag{4.60}$$

Taking into account Lemma 4.4.1, the total energy also simplifies with respect to the one indicated by (4.56):

$$E^s_h(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}) = \frac{1}{2\pi}\int\limits_{\Pi_h} r^s_{ph,h}(\xi)|\widehat{u}^{h,0}(\xi)|^2\,d\xi. \tag{4.61}$$

For initial data with the property (4.54), we consider a second restriction

$$\widehat{u}^{h,1}(\xi) = i\lambda^s_{sp,h}(\xi)\widehat{u}^{h,0}(\xi).$$

With this second restriction, the solution simplifies with respect to the one given by (4.57):

$$\widehat{U}^h(\xi,t) = P^s_{sp,h}(\xi)\widehat{u}^{h,0}(\xi)\exp(it\lambda^s_{sp,h}(\xi)). \tag{4.62}$$

Taking into account Lemma 4.4.1, the total energy also simplifies with respect to the one indicated by (4.58):

$$E^s_h(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1}) = \frac{1}{2\pi}\int_{\Pi_h} r^s_{sp,h}(\xi)|\widehat{u}^{h,0}(\xi)|^2\,d\xi. \tag{4.63}$$

The main result of this section is the following one:

Theorem 4.4.1. *Fix $T>0$ and consider a wave number $\eta_0\in\Pi_h\setminus\{0\}$ and a point $x^*\in(-1,1)$ such that the semi-discrete GCC is not verified, i.e.*

$$|x_h(t)| = |x^* - t\lambda^s_{ph,1}(\eta_0)| < 1,\ \forall t\in[0,T]. \tag{4.64}$$

Consider $\widehat{\sigma}\in C^\infty_c(-1,1)$ and $\gamma=\gamma(h)>0$ to be a function such that the following two condition are verified

$$\lim_{h\to 0}\gamma(h) = +\infty \text{ and } \lim_{h\to 0} h\gamma(h) = 0. \tag{4.65}$$

In the semi-discrete wave equation (4.32) the initial data $(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1})$ are concentrated on the physical mode, i.e. (4.53) is verified

$$\widehat{u}^{h,0}(\xi) = \gamma^{-d/2}\widehat{\sigma}(\gamma^{-1}(\xi-\xi_0))\frac{1}{i\omega_{d,h}(\xi)}\exp(-i\xi\cdot x^*) \text{ and } \widehat{u}^{h,1}(\xi) = i\lambda^s_{ph,h}(\xi)\widehat{u}^{h,0}(\xi). \tag{4.66}$$

Then for all $\alpha>0$, there exists a constant $c^s_\alpha = c^s_\alpha(T,\widehat{\sigma},\eta_0)>0$ not depending on h such that the discrete observability constant $C^s_h(T)$ in (4.18) satisfies $C^s_h(T)\geq c^s_\alpha\gamma^\alpha$.

The method of proof is based on a stationary phase argument like in the proof of Theorem 3.3.1 in Chapter 1. A similar result holds for solutions concentrated on the spurious diagram, i.e. satisfying the requirement (4.54).

4.5 Filtering mechanisms for the SIPG approximation

4.5.1 Concentration on the physical mode + Fourier filtering

For $\delta\in(0,1)$, set $\Pi^\delta_h := [-\pi\delta/h,\pi\delta/h]$. Let us define the space of Fourier filtered data:

$$I^\delta_h = \{\overrightarrow{f}\in\ell^2(h\mathbb{Z}) : \mathrm{supp}(\widehat{f}^h)\subset\Pi^\delta_h\}. \tag{4.67}$$

For $\overrightarrow{f}\in\ell^2(h\mathbb{Z})$, define the projection on I^δ_h of $\overrightarrow{f}$ to be

$$\Gamma^\delta_h f_j = \frac{1}{2\pi}\int_{\Pi^\delta_h}\widehat{f}^h(\xi)\exp(i\xi x_j)\,d\xi. \tag{4.68}$$

We say that the initial data $\overrightarrow{u}^{h,i}$, $i=0,1$, in (4.53) is filtered with parameter $\delta\in(0,1)$ if

$$\overrightarrow{u}^{h,i}\in I^\delta_h,\quad \forall i=0,1. \tag{4.69}$$

Since $\lambda^s_{ph,h}(\xi)$ is increasing in ξ, if the initial data in (4.32) is such that $\overrightarrow{U}^{h,i}\in(I^\delta_h)^2$, $i=0,1$, and such that the condition (4.53) holds, then the maximum frequency involved in the solution (4.47) is $\lambda^s_{ph,h}(\delta\pi/h)$.

The following result of uniform observability holds:

Theorem 4.5.1. *Set $\Omega = \{x : |x| > 1\}$. In (4.15) consider initial data $\overrightarrow{U}^{h,i}$, $i = 0, 1$, concentrated on the physical mode, i.e. verifying condition (4.53), and filtered with parameter $\delta \in (0,1)$, i.e. verifying (4.69). If $T > T_{ph}^{s,\delta}$, with*

$$T_{ph}^{s,\delta} = \frac{2}{\min\limits_{\xi \in \Pi_h^\delta} \partial_\xi \lambda_{ph,h}^s(\xi)}(1 + C_{ph}^{s,\delta})$$

and $C_{ph}^{s,\delta} \in (0,1)$. Then, for all $s > 1$, the following observability inequality holds uniformly as $h \to 0$:

$$E_h^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}) \leq C_{ph}^{s,\delta}(T) \int_0^T E_{\Omega,h}^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\, dt, \tag{4.70}$$

with observability constant

$$C_{ph}^{s,\delta}(T) = \frac{1}{T - T_{ph}^{s,\delta}}.$$

Remark 4.5.1. *We could expect that the minimal time $T_{ph}^{s,\delta}$ should depend also on the mesh size h. The time $T_{ph}^{s,\delta}$ is obtained as the maximum on Π_h^δ of a function depending on ξh. By the change of variable $\xi h = \eta \in \Pi_1^\delta$, that function does not depend anymore on h and therefore $T_{ph}^{s,\delta}$ does not depend on h.*

Remark 4.5.2. *In view of the property (GVPh2) from Proposition 4.3.4, for all $s \in (1,\infty) \setminus \{3\}$, $\lim\limits_{\delta \to 1} T_{ph}^{s,\delta} = \infty$, which is in accordance to the blow-up of the observability constant that we proved in the previous section.*

Since $\lim\limits_{\xi \to 0} |\alpha^+(\xi) - \alpha^-(\xi)| = 0$, we obtain that $T_{ph}^{s,\delta} \to 2$ as $\delta \to 0$, which is the observability time corresponding to the continuous case.

Proof. Consider $I = \{x : |x| \leq 1\}$. For all $T > 0$, in view of the conservation of the total energy, the following identity holds:

$$\int_0^T E_{\Omega,h}^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\, dt = T E_h^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}) - \int_0^T E_{I,h}^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\, dt.$$

In what follows, our aim is to find an upper bound for the energy concentrated in the interior domain I. First of all, observe that

$$\int_0^T E_{I,h}^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\, dt \leq \int_{\mathbb{R}} E_{I,h}^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\, dt = \sum_{i=1}^7 I_i,$$

where

$$I_i = \int_{\mathbb{R}} E_{I,h}^{s,i}(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\, dt,$$

with $E_{I,h}^{s,i}$ defined by (4.26).

Computation of the terms I_i, $i = 1, \cdots, 7$. According to the property (GVPh4) in Proposition 4.3.4, $\lambda_{ph,h}^s(\xi)$ is increasing in Π_h, i.e. is injective. Consequently, the change of variable $\zeta = \lambda_{ph,h}^s(\xi)$ is well defined. Set $\xi(\zeta) = (\lambda_{ph,h}^s)^{-1}(\zeta)$. Observe that

$$d\xi = \frac{1}{\partial_\xi \lambda_{ph,h}^s(\xi(\zeta))}\, d\zeta. \tag{4.71}$$

For all $a, b \in \mathbb{C}$, we have the identity

$$\frac{1}{2}(|a|^2 + |b|^2) = \Big|\frac{a+b}{2}\Big|^2 + \Big|\frac{a-b}{2}\Big|^2. \tag{4.72}$$

Let us denote by

$$\widehat{u}^{h,\pm}(\xi) := \frac{1}{2}\big(i\lambda^s_{ph,h}(\xi)\widehat{u}^{h,0}(\xi) \pm \widehat{u}^{h,1}(\xi)\big).$$

By (4.72), we get that for all $\xi \in \Pi_h$,

$$|\widehat{u}^{h,+}(\xi)|^2 + |\widehat{u}^{h,-}(\xi)|^2 = \frac{1}{2}\big(\Lambda^s_{ph,h}(\xi)|\widehat{u}^{h,0}(\xi)|^2 + |\widehat{u}^{h,1}(\xi)|^2\big). \tag{4.73}$$

Also set

$$\widehat{U}^{h,+}_k(\xi) := \sum_{\pm}[\widehat{u}^{h,\pm}(\pm\xi)\exp(\pm i\xi x_k)], \quad \widehat{U}^{h,-}_k(\xi) := \sum_{\pm}[\pm\widehat{u}^{h,\pm}(\pm\xi)\exp(\pm i\xi x_k)].$$

From (4.72), we get that for all $\xi \in \Pi_h$,

$$|\widehat{U}^{h,+}_k(\xi)|^2 + |\widehat{U}^{h,-}_k(\xi)|^2 = 2\big(|\widehat{u}^{h,+}(\xi)|^2 + |\widehat{u}^{h,-}(-\xi)|^2\big). \tag{4.74}$$

Take into account the fact that both $\lambda^s_{ph,h}(\xi)$ and $p^s_{ph,h}(\xi)$ are odd functions with respect to the wave number $\xi = 0$. Then the solution (4.55) can be reorganized as

$$A_k(t) = \frac{1}{2\pi}\int\limits_{\lambda^s_{ph,h}(\Pi^\delta_h)} \widehat{U}^{h,-}_k(\xi(\zeta))\frac{1}{i\zeta}\frac{1}{\sqrt{1+|p^s_{ph,h}(\xi(\zeta))}}\frac{\exp(it\zeta)}{\partial_\xi\lambda^s_{ph,h}(\xi(\zeta))}\,d\zeta \tag{4.75}$$

and

$$J_k(t) = \frac{1}{2\pi}\int\limits_{\lambda^s_{ph,h}(\Pi^\delta_h)} \widehat{U}^{h,+}_k(\xi(\zeta))\frac{1}{i\zeta}\frac{p^s_{ph,h}(\xi(\zeta))}{\sqrt{1+|p^s_{ph,h}(\xi(\zeta))|^2}}\frac{\exp(it\zeta)}{\partial_\xi\lambda^s_{ph,h}(\xi(\zeta))}\,d\zeta. \tag{4.76}$$

Then

$$\partial_t A_{k+1}(t) + \partial_t A_k(t) = \frac{1}{2\pi}\int\limits_{\lambda^s_{ph,h}(\Pi^\delta_h)} \widehat{U}^{h,-}_{k+1/2}(\xi(\zeta))\frac{2\cos\big(\frac{\xi(\zeta)h}{2}\big)}{\partial_\xi\lambda^s_{ph,h}(\xi(\zeta))}\frac{\exp(it\zeta)}{\sqrt{1+|p^s_{ph,h}(\xi(\zeta))|^2}}\,d\zeta. \tag{4.77}$$

Define

$$d\mu(\xi) := \frac{1}{1+|p^s_{ph,h}(\xi)|^2}\frac{1}{\partial_\xi\lambda^s_{ph,h}(\xi)}\,d\xi.$$

From (4.71), we obtain that

$$d\mu(\xi(\zeta)) := \frac{1}{1+|p^s_{ph,h}(\xi(\zeta))|^2}\frac{1}{|\partial_\xi\lambda^s_{ph,h}(\xi(\zeta))|^2}\,d\zeta.$$

Observe that the integral in the right hand side of (4.77) is in fact a continuous inverse Fourier transform in t. Then, by Parseval identity, we get

$$I_1 = \frac{h}{24}\sum_{|x_k|<1}\frac{1}{2\pi}\int\limits_{\lambda^s_{ph,h}(\Pi^\delta_h)}\Big[|\widehat{U}^{h,-}_{k+1/2}(\xi(\zeta))|^2 + |\widehat{U}^{h,-}_{k-1/2}(\xi(\zeta))|^2\Big]4\cos^2\big(\frac{\xi(\zeta)h}{2}\big)\,d\mu(\xi(\zeta)). \tag{4.78}$$

Using (4.72), we obtain

$$\frac{1}{2}[|\widehat{U}^{h,-}_{k+1/2}(\xi)|^2 + |\widehat{U}^{h,-}_{k-1/2}(\xi)|^2] = |\widehat{U}^{h,-}_k(\xi)|^2\cos^2\big(\frac{\xi h}{2}\big) + |\widehat{U}^{h,+}_k(\xi)|^2\sin^2\big(\frac{\xi h}{2}\big). \tag{4.79}$$

Undoing the change of variable $\zeta = \lambda^s_{ph,h}(\xi)$ in (4.78) and applying (4.79), we obtain

$$I_1 = \frac{h}{12} \sum_{|x_k|<1} \frac{1}{2\pi} \int\limits_{\Pi_h^\delta} \Big[4\cos^4\big(\frac{\xi h}{2}\big) |\widehat{U}_k^{h,-}(\xi)|^2 + \sin^2(\xi h) |\widehat{U}_k^{h,+}(\xi)|^2 \Big] \, d\mu(\xi). \tag{4.80}$$

Using the same arguments as for I_1, we obtain

$$I_2 = \frac{h}{2} \sum_{|x_k|<1} \frac{1}{2\pi} \int\limits_{\Pi_h^\delta} \Big[\cos^2\big(\frac{\xi h}{2}\big) |\widehat{U}_k^{h,+}(\xi)|^2 + \sin^2\big(\frac{\xi h}{2}\big) |\widehat{U}_k^{h,-}(\xi)|^2 \Big] \frac{\Omega_h(\xi)}{\Lambda^s_{ph,h}(\xi)} \, d\mu(\xi), \tag{4.81}$$

$$I_3 = \frac{h}{48} \sum_{|x_k|<1} \frac{1}{2\pi} \int\limits_{\Pi_h^\delta} \Big[\sin^2(\xi h) |\widehat{U}_k^{h,-}(\xi)|^2 + 4\sin^4\big(\frac{\xi h}{2}\big) |\widehat{U}_k^{h,+}(\xi)|^2 \Big] |p^s_{ph,h}(\xi)|^2 \, d\mu(\xi), \tag{4.82}$$

$$I_4 = \frac{h}{8} \sum_{|x_k|<1} \frac{1}{2\pi} \int\limits_{\Pi_h^\delta} \Big[\cos^2\big(\frac{\xi h}{2}\big) |\widehat{U}_k^{h,-}(\xi)|^2 + \sin^2\big(\frac{\xi h}{2}\big) |\widehat{U}_k^{h,+}(\xi)|^2 \Big] \frac{\Omega_h(\xi)}{\Lambda^s_{ph,h}(\xi)} |p^s_{ph,h}(\xi)|^2 \, d\mu(\xi) \tag{4.83}$$

and

$$I_5 = \frac{s-1}{2h} \sum_{|x_k|<1} \frac{1}{2\pi} \int\limits_{\Pi_h^\delta} |\widehat{U}_k^{h,+}(\xi)|^2 \frac{|p^s_{ph,h}(\xi)|^2}{\Lambda^s_{ph,h}(\xi)} \, d\mu(\xi). \tag{4.84}$$

By applying the same arguments as for I_1, we obtain that I_6 and I_7 are given explicitly by

$$I_6 = \frac{h}{24} \sum_{|x_k|<1} \frac{1}{2\pi} \int\limits_{\Pi_h^\delta} (a_k^1(\xi) + a_k^2(\xi)) \, d\mu(\xi) \tag{4.85}$$

and

$$I_7 = \frac{h}{24} \sum_{|x_k|<1} \frac{1}{2\pi} \int\limits_{\Pi_h^\delta} (a_k^3(\xi) + a_k^4(\xi)) \, d\mu(\xi), \tag{4.86}$$

where

$$a_k^1(\xi) = \Big| \sum_{\pm} \big[\pm \widehat{u}^{h,\pm}(\pm\xi) \exp(\pm i\xi x_k) \big(1 \pm \frac{1}{2} \exp(\pm i\xi h) p^s_{ph,h}(\xi)\big) \big] \Big|^2,$$

$$a_k^2(\xi) = \Big| \sum_{\pm} \big[\pm \widehat{u}^{h,\pm}(\pm\xi) \exp(\pm i\xi x_k) \big(\exp(\mp i\xi h) \pm \frac{1}{2} p^s_{ph,h}(\xi)\big) \big] \Big|^2,$$

$$a_k^3(\xi) = \Big| \sum_{\pm} \big[\pm \widehat{u}^{h,\pm}(\pm\xi) \exp(\pm i\xi x_k) \big(\exp(\pm i\xi h) \mp \frac{1}{2} p^s_{ph,h}(\xi)\big) \big] \Big|^2$$

and

$$a_k^4(\xi) = \Big| \sum_{\pm} \big[\pm \widehat{u}^{h,\pm}(\pm\xi) \exp(\pm i\xi x_k) \big(1 \mp \frac{1}{2} p^s_{ph,h}(\xi) \exp(\mp i\xi h)\big) \big] \Big|^2.$$

By applying formula (4.72), we obtain

$$\frac{1}{2}(a_k^2(\xi) + a_k^3(\xi)) = |\widehat{U}_k^{h,-}(\xi)|^2 \cos^2(\xi h) + |\widehat{U}_k^{h,+}(\xi)|^2 \big| \sin(\xi h) + \frac{i}{2} p^s_{ph,h}(\xi) \big|^2$$

and

$$\frac{1}{2}(a_k^1(\xi) + a_k^4(\xi)) = |\widehat{U}_k^{h,-}(\xi)|^2 \big| 1 + \frac{i}{2} \sin(\xi h) p^s_{ph,h}(\xi) \big|^2 + |\widehat{U}_k^{h,+}(\xi)|^2 \frac{1}{4} |p^s_{ph,h}(\xi)|^2 \cos^2(\xi h).$$

Then

$$I_6 + I_7 = \frac{h}{12} \sum_{|x_k|<1} \int_{\Pi_h^\delta} \Big[\big(\cos^2(\xi h) + \big| 1 + \frac{i}{2} \sin(\xi h) p^s_{ph,h}(\xi) \big|^2 \big) |\widehat{U}^{h,-}_k(\xi)|^2 \tag{4.87}$$
$$+ \big(\big| \sin(\xi h) + \frac{i}{2} p^s_{ph,h}(\xi) \big|^2 + \frac{1}{4} |p^s_{ph,h}(\xi)|^2 \cos^2(\xi h) \big) |\widehat{U}^{h,+}_k(\xi)|^2 \Big] \, d\mu(\xi).$$

In view of the identities (4.80-4.84) and (4.87), we conclude that

$$\int_0^T E^s_{I,h}(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t) \, dt \leq \frac{h}{2} \sum_{|x_k|<1} \frac{1}{2\pi} \int_{\Pi_h^\delta} \left[\alpha^+(\xi) |\widehat{U}^{h,+}_k(\xi)|^2 + \alpha^-(\xi) |\widehat{U}^{h,-}_k(\xi)|^2 \right] d\mu(\xi), \tag{4.88}$$

where

$$\begin{aligned} \alpha^+(\xi) = {} & \frac{1}{6} \sin^2(\xi h) + \cos^2\big(\frac{\xi h}{2}\big) \frac{\Omega_h(\xi)}{\Lambda^s_{ph,h}(\xi)} + \frac{1}{6} \sin^4\big(\frac{\xi h}{2}\big) |p^s_{ph,h}(\xi)|^2 \\ & + \frac{1}{4} \sin^2\big(\frac{\xi h}{2}\big) \frac{\Omega_h(\xi)}{\Lambda^s_{ph,h}(\xi)} |p^s_{ph,h}(\xi)|^2 \\ & + \frac{s-1}{h^2} \frac{|p^s_{ph,h}(\xi)|^2}{\Lambda^s_{ph,h}(\xi)} + \frac{1}{6} \Big(\Big| \sin(\xi h) + \frac{i}{2} p^s_{ph,h}(\xi) \Big|^2 + \frac{1}{4} |p^s_{ph,h}(\xi)|^2 \cos^2(\xi h) \Big) \end{aligned}$$

and

$$\begin{aligned} \alpha^-(\xi) = {} & \frac{2}{3} \cos^4\big(\frac{\xi h}{2}\big) + \sin^2\big(\frac{\xi h}{2}\big) \frac{\Omega_h(\xi)}{\Lambda^s_{ph,h}(\xi)} + \frac{1}{24} \sin^2(\xi h) |p^s_{ph,h}(\xi)|^2 \\ & + \frac{1}{4} \cos^2\big(\frac{\xi h}{2}\big) \frac{\Omega_h(\xi)}{\Lambda^s_{ph,h}(\xi)} |p^s_{ph,h}(\xi)|^2 \\ & + \frac{1}{6} \Big(\cos^2(\xi h) + \big| 1 + \frac{i}{2} \sin(\xi h) p^s_{ph,h}(\xi) \big|^2 \Big). \end{aligned} \tag{4.89}$$

In what follows, we will use the following lemma:

Lemma 4.5.1. *For all $\xi \in \Pi_h$, the following identity holds:*

$$\frac{\alpha^+(\xi) + \alpha^-(\xi)}{2} = m^s_{ph,h}(\xi) \big(1 + |p^s_{ph,h}(\xi)|^2 \big).$$

The proof of this lemma can be done by identifying the the expressions of $m^s_{ph,h}(\xi)$ and $r^s_{ph,h}(\xi)$ introduced in (4.56) in the expression of $\alpha^+(\xi) + \alpha^-(\xi)$ and by using Lemma 4.4.1.

Using the above lemma and (4.74), for all $k \in \mathbb{Z}$ we have

$$\begin{aligned} & |\widehat{U}^{h,+}_k(\xi)|^2 \alpha^+(\xi) + |\widehat{U}^{h,-}_k(\xi)|^2 \alpha^-(\xi) \\ & \qquad \leq \Big(\frac{\alpha^+(\xi) + \alpha^-(\xi)}{2} + \frac{|\alpha^+(\xi) - \alpha^-(\xi)|}{2} \Big) \big(|\widehat{U}^{h,+}_k(\xi)|^2 + |\widehat{U}^{h,-}_k(\xi)|^2 \big) \\ & \qquad = 2(1 + C^s_{ph,h}(\xi)) m^s_{ph,h}(\xi) \big(1 + |p^s_{ph,h}(\xi)|^2 \big) \big(|\widehat{u}^{h,+}(\xi)|^2 + |\widehat{u}^{h,-}(-\xi)|^2 \big), \end{aligned}$$

where

$$C^s_{ph,h}(\xi) = \frac{|\alpha^+(\xi) - \alpha^-(\xi)|}{2 m^s_{ph,h}(\xi)(1 + |p^s_{ph,h}(\xi)|^2)}.$$

Using Lemma 4.5.1, we have that $C^s_{ph,h}(\xi) \in [0, 1]$. Observe that within $C^s_{ph,h}(\xi)$ is in fact a rational function involving only trigonometric functions depending on $\eta = \xi h \in \Pi_1$. Then

$$C^s_{ph,h}(\xi) = C^s_{ph,1}(\eta).$$

Taking into account the fact that the right hand side of the above inequality does not depend on k and that both $m^s_{ph,h}(\xi)$ and $C^s_{ph,h}(\xi)$ are even functions in ξ, we obtain

$$\int_0^T E^s_{I,h}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)\,dt \leq \frac{1}{2\pi}\int_{\Pi_h^\delta} 2(1+C^s_{ph,h}(\xi))m^s_{ph,h}(\xi)(|\widehat{u}^{h,+}(\xi)|^2+|\widehat{u}^{h,-}(\xi)|^2)\frac{1}{\partial_\xi\lambda^s_{ph,h}(\xi)}\,d\xi.$$

By Lemma 4.4.1 and (4.73), we get

$$\int_0^T E^s_{I,h}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)\,dt \leq \frac{1}{2\pi}\int_{\Pi_h^\delta} \frac{1+C^s_{ph,h}(\xi)}{\partial_\xi\lambda^s_{ph,h}(\xi)}\big(m^s_{ph,h}(\xi)|\widehat{u}^{h,1}(\xi)+r^s_{ph,h}(\xi)|\widehat{u}^{h,0}(\xi)|^2\big)\,d\xi.$$

The conclusion of the theorem follows from the fact that since $C^s_{ph,h}(\xi)\in L^\infty(\Pi_h)$, there exists a constant

$$C^{s,\delta}_{ph}=\max_{\xi\in\Pi_h^\delta} C^s_{ph,h}(\xi)=\max_{\eta\in\Pi_1^\delta} C^s_{ph,1}(\eta).$$

□

4.5.2 Concentration on the physical mode + bi-grid algorithm

In this section, we consider initial data $\overrightarrow{u}^{h,i}$, $i=0,1$, in (4.53) given by a bi-grid algorithm. More precisely, for $i=0,1$, the initial data $\overrightarrow{u}^{h,i}$ is obtained by linear interpolation from a coarse grid of size $2h$, G^{2h}, to a fine one of size h, G^h. Explicitly, the odd components are related to the even ones in the following way:

$$u^{h,i}_{2j}=\frac{u^{h,i}_{2j-1}+u^{h,i}_{2j+1}}{2}, i=0,1. \tag{4.90}$$

The main result of this subsection is the following one.

Theorem 4.5.2. *Set $\Omega=\{x:|x|>1\}$. In (4.15) consider initial data $\overrightarrow{U}^{h,i}$, $i=0,1$, concentrated on the physical mode, i.e. verifying the condition (4.53), such that, for $i=0,1$, $\overrightarrow{u}^{h,i}$ satisfies the bi-grid condition (4.90). For all $T>T^{s,1/2}_{ph}$, with $T^{s,\delta}_{ph}$ given in Theorem 4.5.1 and all $s>1$, there exists a constant $C_{ph,bigrid}(T,s)>0$ independent of h such that the following observability inequality holds:*

$$E^s_h(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1})\leq C_{ph,bigrid}(T,s)\int_0^T E^s_{\Omega,h}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)\,dt. \tag{4.91}$$

In Figure 4.8, we represent the two components of the solution of the SIPG semi-discretization of the wave equation: $\overrightarrow{A}(t)$ (green) and $\overrightarrow{J}(t)$ (red) for $s=5$, at time $t=1$, starting with $\widehat{U}^{h,0}(\xi)=\sqrt{2\pi/\gamma^2}\exp(-(\xi-\pi/h)^2/(2\gamma^2))P^s_{ph,h}(\xi)$, $\widehat{U}^{h,1}(\xi)=i\lambda^s_{ph,h}(\xi)\widehat{U}^{h,0}(\xi)$, $h=0.001$ and $\gamma=h^{-2/3}$, without bigrid (top), with bi-grid of ratio 1/2 (middle) and with bigrid of ratio 1/4 (bottom).

Proof of Theorem 4.5.2. Like in the Proof of Theorem 6.2.1., pp. 159, in [30], or [34], we will apply a dyadic decomposition argument which has to be accompanied by the following result whose proof will be provided after the proof of this theorem:

Proposition 4.5.1. *For all $\overrightarrow{U}^{h,i}$ satisfying (4.53), $\overrightarrow{u}^{h,i}$ verifying (4.90), $i=0,1$, and $s\in(1,\infty)$, there exists a constant $C_{ph}>0$, independent of h and of s, such that*

$$E^s_h(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1})\leq C_{ph}E^s_h(\Gamma_h^{1/2}\overrightarrow{U}^{h,0},\Gamma_h^{1/2}\overrightarrow{U}^{h,1}), \tag{4.92}$$

where Γ_h^δ is defined by (4.68).

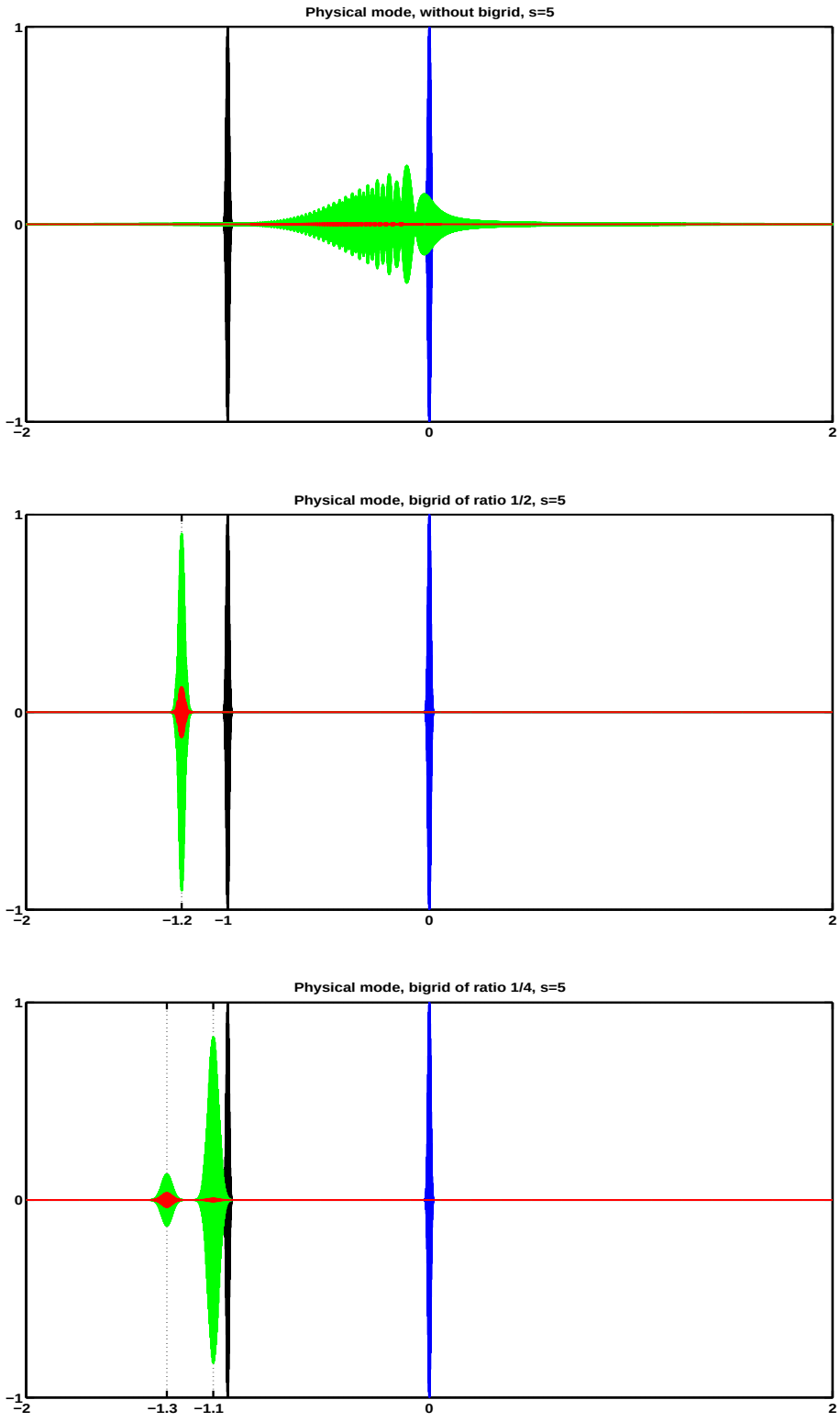

Figure 4.8: Propagation of waves packets for the SIPG semi-discrete wave equations with different types of bi-grids.

The dyadic decomposition technique was also used in [32] in order to prove the boundary observability of the classical finite difference semi-discretization of the $2-d$ wave equation in the unit square when the initial data is given by bi-grid algorithm.

We divide the proof of this theorem into several steps as follows:

Step 1: Consequences of the continuity of the minimal time $T_{ph}^{s,\delta}$ with respect to δ. Let T be an observability time provided by Theorem 4.5.1 under the assumptions that the initial data $\overrightarrow{U}^{h,i}$ are concentrated only on the physical mode, i.e. they verify (4.53) and $\overrightarrow{u}^{h,i} \in I_h^\delta$. By the continuous dependence of $T_{ph}^{s,\delta}$ with respect to δ, there exist $\gamma, \epsilon > 0$ such that $T - 4\gamma > T_{ph}^{s,\delta+\epsilon} > T_{ph}^{s,\delta}$ and the following observability inequality holds uniformly as $h \to 0$:

$$E_h^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}) \leq C_{ph}^{s,\delta+\epsilon}(T-4\gamma) \int_{2\gamma}^{T-2\gamma} E_{\Omega,h}^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\, dt, \tag{4.93}$$

for all $\overrightarrow{U}^{h,i}$ verifying (4.53) and $\overrightarrow{u}^{h,i} \in I_h^{\delta+\epsilon}$, $i = 0, 1$.

Step 2: Definition and total energy for the projectors. Denote by $\widehat{\cdot}$ the continuous Fourier transform. For $c > 1$ and $P \in C_c^\infty(\mathbb{R})$ and $f \in L^1(\mathbb{R})$, we introduce the projector $P_k f$ in the following way:

$$(P_k f)(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(\tau) P\Big(\frac{\tau}{c^k}\Big) \exp(it\tau)\, d\tau. \tag{4.94}$$

For a, b, c, μ verifying the following conditions

$$1 < c < \frac{b-\mu}{a+\mu} < \frac{b}{a} < \frac{\lambda_{ph,h}^s((1/2+\epsilon)\pi/h)}{\lambda_{ph,h}^s(\pi/2h)}, \tag{4.95}$$

we consider

$$F \in C_c^\infty(a,b), \quad 0 \leq F \leq 1, \quad F \equiv 1 \text{ in } (a+\mu, b-\mu) \tag{4.96}$$

and

$$P(\tau) = \begin{cases} F(\tau), & \tau > 0 \\ F(-\tau), & \tau < 0. \end{cases} \tag{4.97}$$

With this P, we construct the projectors P_k defined by (4.94).

In the particular case when the initial data in (4.15) satisfy both conditions (4.53) and (4.90), taking into account the form of the solution (4.55), the projectors $P_k U_j(t)$ are given by

$$P_k U_j(t) = \frac{1}{2\pi} \int_{\Pi_h} P\big(\lambda_{ph,h}^s(\xi)/c^k\big) P_{ph,h}^s(\xi) \widehat{u}^h(\xi, t) \exp(i\xi x_j)\, d\xi, \tag{4.98}$$

where

$$\widehat{u}^h(\xi, t) = \frac{1}{2} \sum_{\pm} \Big(\widehat{u}^{h,0}(\xi) \pm \frac{\widehat{u}^{h,1}(\xi)}{i\lambda_{ph,h}^s(\xi)} \Big) \exp(\pm it\lambda_{ph,h}^s(\xi)).$$

For all $k \in \mathbb{N}$, $P_k \overrightarrow{U}^h(t)$ is a solution of (4.15) corresponding to initial data $\overrightarrow{V}^{h,i}$ with

$$\widehat{V}^{h,i}(\xi) = P_{ph,h}^s(\xi) P\Big(\frac{\lambda_{ph,h}^s(\xi)}{c^k}\Big) \widehat{u}^{h,i}(\xi). \tag{4.99}$$

Then the total energy associated to $P_k \overrightarrow{U}^h(t)$, $E_h^s(P_k \overrightarrow{U}^h(t), \partial_t P_k \overrightarrow{U}^h(t)$, is conserved in time. Denote it by $E_h^s(P_k \overrightarrow{U}^{h,0}, P_k \overrightarrow{U}^{h,1})$. On the other hand, $\widehat{V}^{h,i}(\xi)$ verifies (4.53). Then

$$E_h^s(P_k \overrightarrow{U}^{h,0}, P_k \overrightarrow{U}^{h,1}) = \frac{1}{4\pi} \int_{\Pi_h} P^2\Big(\frac{\lambda_{ph,h}^s(\xi)}{c^k}\Big) \big[m_{ph,h}^s(\xi) |\widehat{u}^{h,1}(\xi)|^2 + r_{ph,h}^s(\xi) |\widehat{u}^{h,0}(\xi)|^2 \big]\, d\xi. \tag{4.100}$$

Step 3. Apply Proposition 4.5.1.

Step 4. Choice of the upper bound for the index k of the projectors P_k. Since for all $k_0 \in \mathbb{N}$, $\bigcup_{k=k_0}^{\infty} (ac^k, bc^k) = (ac^{k_0}, \infty)$, deduce that any $\lambda^s_{ph,h}(\xi) > ac^{k_0}$ is located at least in the support of one projector P_k, $k \geq k_0$. For any $h > 0$ and $\delta = 1/2$ consider the unique k_h s.t.

$$c^{k_h}(a+\mu) \leq \lambda^s_{ph,h}(\pi\delta/h) < c^{k_h+1}(a+\mu). \tag{4.101}$$

By making use of the condition (4.95), with $\delta = 1/2$ and $k_0 \in \mathbb{N}$ independent of h and $k = k_0, \cdots, k_h$, we have

$$(a+\mu)c^{k_0} \leq c^k(a+\mu) \leq c^{k_h}(a+\mu) \leq \lambda^s_{ph,h}(\pi\delta/h) < c^{k_h+1}(a+\mu) < c^{k_h}(b-\mu). \tag{4.102}$$

Then, for any $k_0 \in \mathbb{N}$ not depending on h, we deduce that any frequency $\lambda^s_{ph,h}(\xi) \in ((a+\mu)c^{k_0}, \lambda^s_{ph,h}(\pi/2h))$ is contained in at least one interval $((a+\mu)c^k, (b-\mu)c^k)$, $k_0 \leq k \leq k_h$, i.e. in the region where $P(\cdot/c^k) \equiv 1$ and, consequently, for $\delta = 1/2$,

$$1 \leq \sum_{k=k_0}^{k_h} P\Big(\frac{\lambda^s_{ph,h}(\xi)}{c^k}\Big), \ \forall \xi \text{ s.t. } \lambda^s_{ph,h}(\xi) \in ((a+\mu)c^{k_0}, \lambda^s_{ph,h}(\pi\delta/h)). \tag{4.103}$$

We will choose k_0 later on.

Step 5. Upper bounds of $E^s_h(\Gamma_h^{1/2}\overrightarrow{U}^{h,0}, \Gamma_h^{1/2}\overrightarrow{U}^{h,1})$ in terms of the energy of the projections, $E^s_h(P_k\overrightarrow{U}^{h,0}, P_k\overrightarrow{U}^{h,1})$. Consider

$$\Pi_h^{k_0} = \{\xi \in \Pi_h : |\lambda^s_{ph,h}(\xi)| \leq c^{k_0}(a+\mu)\}. \tag{4.104}$$

For $\overrightarrow{f} \in \ell^2(h\mathbb{Z})$, define its projection on $\Pi_h^{k_0}$ to be

$$\Gamma_{k_0} f_j = \frac{1}{2\pi} \int\limits_{\Pi_h^{k_0}} \widehat{f^h}(\xi) \exp(i\xi x_j)\, d\xi. \tag{4.105}$$

From (4.68) and (4.105), we see that, for any $\delta \in (0,1)$,

$$\big(\Gamma_h^\delta - \Gamma_{k_0}\big) f_j = \frac{1}{2\pi} \int\limits_{\Pi_h^\delta \setminus \Pi_h^{k_0}} \widehat{f^h}(\xi) \exp(i\xi x_j)\, d\xi.$$

Consequently, the energy $E^s_h(\Gamma_h^{1/2}\overrightarrow{U}^{h,0}, \Gamma_h^{1/2}\overrightarrow{U}^{h,1})$ can be decomposed as follows

$$\begin{aligned} E^s_h(\Gamma_h^{1/2}\overrightarrow{U}^{h,0}, &\Gamma_h^{1/2}\overrightarrow{U}^{h,1}) = \\ &= E^s_h(\Gamma_{k_0}\overrightarrow{U}^{h,0}, \Gamma_{k_0}\overrightarrow{U}^{h,1}) + E^s_h((\Gamma_h^{1/2} - \Gamma_{k_0})\overrightarrow{U}^{h,0}, (\Gamma_h^{1/2} - \Gamma_{k_0})\overrightarrow{U}^{h,1}). \end{aligned} \tag{4.106}$$

Using (4.103) and (4.100), we get

$$E^s_h((\Gamma_h^{1/2} - \Gamma_{k_0})\overrightarrow{U}^{h,0}, (\Gamma_h^{1/2} - \Gamma_{k_0})\overrightarrow{U}^{h,1}) \leq \sum_{k=k_0}^{k_h} E^s_h(P_k\overrightarrow{U}^{h,0}, P_k\overrightarrow{U}^{h,1}). \tag{4.107}$$

Therefore, from (4.106) and (4.107) we obtain

$$E^s_h(\Gamma_h^{1/2}\overrightarrow{U}^{h,0}, \Gamma_h^{1/2}\overrightarrow{U}^{h,1}) \leq E^s_h(\Gamma_{k_0}\overrightarrow{U}^{h,0}, \Gamma_{k_0}\overrightarrow{U}^{h,1}) + \sum_{k=k_0}^{k_h} E^s_h(P_k\overrightarrow{U}^{h,0}, P_k\overrightarrow{U}^{h,1}). \tag{4.108}$$

The first term in the right hand side of (4.108) will be proved to be a lower order one and will be absorbed by compactness.

Step 6. $P_k\overrightarrow{U}^{h,i} \in I_h^{1/2+\epsilon}$**.** Remark that for all $k_0 \leq k \leq k_h$, the projector $P_k\overrightarrow{U}^{h,i}$, $i=0,1$, contains only frequencies $\lambda^s_{ph,h}(\xi) \in (ac^k, bc^k)$, its support. Taking into account (4.95) and (4.101), we have that for all $k \leq k_h$, the frequencies $\lambda^s_{ph,h}(\xi)$ involved in $P_k\overrightarrow{U}^{h,i}$ satisfy

$$\lambda^s_{ph,h}(\xi) \leq c^k b < c^{k_h} b \leq \lambda^s_{ph,h}(\pi/2h)\frac{b}{(a+\mu)} \leq \lambda^s_{ph,h}((1/2+\epsilon)\pi/h). \tag{4.109}$$

Taking into account the fact that $\lambda^s_{ph,h}(\xi)$ is increasing in ξ, we obtain that $|\xi| \leq (1/2+\epsilon)\pi/h$. This implies that $P_k\overrightarrow{U}^{h,i} \in (I_h^{1/2+\epsilon})^2$, $i=0,1$, for all $k_0 \leq k \leq k_h$.

Step 7. Observability inequality for the projections $P_k\overrightarrow{U}^h(t)$**.** Using (4.93) and (4.109), for all $T-4\gamma > T_{ph}^{s,1/2+\epsilon}$, the following observability inequality holds for each $k_0 \leq k \leq k_h$, with $\delta = 1/2$:

$$E_h^s(P_k\overrightarrow{U}^{h,0}, P_k\overrightarrow{U}^{h,1}) \leq C_{ph}^{s,\delta+\epsilon}(T-4\gamma)\int_{2\gamma}^{T-2\gamma} E^s_{\Omega,h}(P_k\overrightarrow{U}^{h,0}, P_k\overrightarrow{U}^{h,0}, t)\,dt. \tag{4.110}$$

Step 8. Observability inequality for the solution $\overrightarrow{U}^h(t)$**.** We use the following Lemma:

Lemma 4.5.2. *Let μ be a Borel measure, Ω a μ-measurable set such that $\mu(\Omega) < \infty$, P_k the projectors defined by (4.94) and $P \in C_c^\infty(\mathbb{R})$ the function generating the projectors P_k. We consider X to be a Hilbert space and $\|\cdot\|_X$ the associated norm. For any positive T, $\gamma < T/4$ and $c > 1$, there are positive constants $C(P,c)$ and $C(P,T,\gamma)$ such that the following inequality holds*

$$\sum_{k\geqslant k_0}\int_{2\gamma}^{T-2\gamma} \|P_k w(t)\|_X^2\,dt \leq C(P,c)\int_0^T \|w(t)\|_X^2\,dt + \frac{C(P,T,\gamma)}{c^{2k_0}}\sup_{j\in\mathbb{Z}}\|w\|^2_{L^2(lT,(l+1)T,X)}, \tag{4.111}$$

for all positive integers k_0 and $w \in L^2(\mathbb{R},X)$.

For further details concerning the proof of Lemma 4.5.2, see [32], Appendix A. We also take into account the direct inequality

$$\int_{lT}^{(l+1)T} E^s_{\Omega,h}(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\,dt \leq TE_h^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}). \tag{4.112}$$

Step 9. Choice of k_0**.** Let us consider k_0 a positive integer such that

$$C_{ph}TC_{ph}^{s,1/2+\epsilon}(T-4\gamma)C(P,T,\gamma)/c^{2k_0} \leq 1/2, \tag{4.113}$$

where C_{ph} is the constant introduced in Proposition 4.5.1.

We apply Lemma 4.5.2 with

$$X = \{(\overrightarrow{f}^{h,0}, \overrightarrow{f}^{h,1}) : \|(\overrightarrow{f}^{h,0}, \overrightarrow{f}^{h,1})(t)\|_X^2 = E^s_{\Omega,h}(\overrightarrow{f}^{h,0}, \overrightarrow{f}^{h,1}, t) < \infty, \forall t \in (0,T)\}.$$

Therefore, from (4.110), Lemma 4.5.2 and the choice (4.113) of k_0, we have:

$$\sum_{k=k_0}^{k_h} E_h^s(P_k\overrightarrow{U}^{h,0}, P_k\overrightarrow{U}^{h,0}) \leq C_{ph}^{s,1/2+\epsilon}(T-4\gamma)\sum_{k=k_0}^{k_h}\int_{2\gamma}^{T-2\gamma} E^s_{\Omega,h}(P_k\overrightarrow{U}^{h,0}, P_k\overrightarrow{U}^{h,1}, t)\,dt \tag{4.114}$$

$$\leq C_{ph}^{s,1/2+\epsilon}(T-4\gamma)C(P,c)\int_0^T E^s_{\Omega,h}(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\,dt + \frac{1}{2C_{ph}}E_h^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}).$$

In conclusion, combining (4.92), (4.108) and (4.110), we have

$$E_h^s(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1}) \leq \widetilde{C}(T,s,\epsilon,\gamma,P,c)\int_0^T E_{\Omega,h}^s(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)\,dt + 2C_{ph}E_h^s(\Gamma_{k_0}\overrightarrow{U}^{h,0},\Gamma_{k_0}\overrightarrow{U}^{h,1}), \tag{4.115}$$

where $\widetilde{C}(T,s,\epsilon,\gamma,P,c) = 2C_{ph}C_{ph}^{s,1/2+\epsilon}(T-4\gamma)C(P,c)$.

Step 10. Compactness of the term $E_h^s(\Gamma_{k_0}\overrightarrow{U}^{h,0},\Gamma_{k_0}\overrightarrow{U}^{h,1})$.
This concludes the proof of Theorem 4.5.2. □

In order to prove Proposition 4.5.1, we will apply the following well-known result:

Lemma 4.5.3. *If $\overrightarrow{f}^h = (f_j)_{j\in\mathbb{Z}}$ satisfies*

$$f_{2j} = \frac{f_{2j+1}+f_{2j-1}}{2},$$

then $\widehat{f}^h(\xi)$ is a π/h-periodic function. Moreover, for any $\xi\in\Pi_h$, $\widehat{f}^h(\xi)$ and the SDFT at scale $2h$ of $(f_{2j+1})_{j\in\mathbb{Z}}$, $\widehat{f}^{2h}(\xi)$, are related as follows

$$\widehat{f}^h(\xi) = \cos^2\Big(\frac{\xi h}{2}\Big)\widehat{f}^{2h}(\xi). \tag{4.116}$$

For the proof of the above Lemma, see [30], Lemma 6.4.1., pp. 170.

Proof of Proposition 4.5.1. Form the fact that the initial data $\overrightarrow{u}^{h,i}$, $i=0,1$, verify (4.90), by Lemma 4.5.3, we have

$$\widehat{u}^{h,i}(\xi) = \cos^2\Big(\frac{\xi h}{2}\Big)\widehat{u}^{2h,i}(\xi),\ \xi\in\Pi_h. \tag{4.117}$$

Taking into account (4.53), the total energy $E_h^s(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1})$ has the simpler form (4.56). The following decomposition of the energy holds

$$E_h^s(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1}) = E_h^s(\Gamma_h^{1/2}\overrightarrow{U}^{h,0},\Gamma_h^{1/2}\overrightarrow{U}^{h,1}) + E_h^s((\Gamma_h^1-\Gamma_h^{1/2})\overrightarrow{U}^{h,0},(\Gamma_h^1-\Gamma_h^{1/2})\overrightarrow{U}^{h,1}). \tag{4.118}$$

Taking into account (4.117) and the π/h-periodicity of $\widehat{u}^{h,i}$, $i=0,1$, we have

$$\begin{aligned}&E_h^s((\Gamma_h^1-\Gamma_h^{1/2})\overrightarrow{U}^{h,0},(\Gamma_h^1-\Gamma_h^{1/2})\overrightarrow{U}^{h,1}) = \\ &\quad = \frac{1}{4\pi}\int_{\Pi_{2h}}\Big(m_{ph,h}^s\big(\xi-\mathrm{sign}(\xi)\pi/h\big)\widehat{u}^{2h,1}(\xi) + r_{ph,h}^s\big(\xi-\mathrm{sign}(\xi)\pi/h\big)\widehat{u}^{2h,0}(\xi)\Big)\sin^4\Big(\frac{\xi h}{2}\Big)\,d\xi.\end{aligned}$$

From the above expression of $E_h^s((\Gamma_h^1-\Gamma_h^{1/2})\overrightarrow{U}^{h,0},(\Gamma_h^1-\Gamma_h^{1/2})\overrightarrow{U}^{h,1})$ as integral on Π_{2h}, observe that it is sufficient to show that both quantities $C_{ph,h}^{s,m}$ and $C_{ph,h}^{s,r}$ defined by

$$C_{ph,h}^{s,m} = \max_{\xi\in\Pi_{2h}} C_{ph,h}^{s,m}(\xi) = \max_{\xi\in\Pi_{2h}}\frac{m_{ph,h}^s(\xi\pm\pi/h)}{m_{ph}^s(\xi)}\tan^4\Big(\frac{\xi h}{2}\Big)$$

and

$$C_{ph,h}^{s,r} = \max_{\xi\in\Pi_{2h}} C_{ph,h}^{s,r}(\xi) = \max_{\xi\in\Pi_{2h}}\frac{r_{ph,h}^s(\xi\pm\pi/h)}{r_{ph}^s(\xi)}\tan^4\Big(\frac{\xi h}{2}\Big)$$

exist and do not depend on h and s. After proving these two things, we take $C_{ph} = 1+\max\{C_{ph,h}^{s,m},C_{ph,h}^{s,r}\}$.

Firstly, let us observe that taking into account the explicit form of $m^s_{ph,h}$ introduced in (4.56) and Proposition 4.3.2, we have that $C^{s,m}_{ph,h}(\xi)$ could be singular as $\xi \to 0$ when $s \in (1,3)$. We have

$$\lim_{\xi\to 0} m^s_{ph,h}(\xi \pm \pi/h) = \begin{cases} 1/4, & s \in (1,3) \\ 1/3, & s \in (3,\infty) \end{cases}$$

and $\lim_{\xi\to 0} m^s_{ph,h}(\xi) = 1$. From these limits, we conclude that for all $s \in (1,\infty)\setminus\{3\}$, we have $\lim_{\xi\to 0} C^{s,m}_{ph,h}(\xi) = 0$. Also, since $\lim_{s\to\infty} p^s_{ph,h}(\xi) = 0$, we have

$$\lim_{s\to\infty} C^{s,m}_{ph,h}(\xi) = \frac{2-\cos(\xi h)}{2+\cos(\xi h)} \tan^4\Big(\frac{\xi h}{2}\Big),$$

which is finite in ξ. Then there exists $C^{s,m}_{ph,h}$ uniformly as $s \to \infty$. The independency with respect to h follows by rescaling, from the fact that

$$C^{s,m}_{ph,h} = \max_{\xi\in\Pi_{2h}} C^{s,m}_{ph,h}(\xi) = \max_{\xi\in\Pi_2} C^{s,m}_{ph,1}(\xi),$$

which does not depend on h.

Analogously, taking into account the explicit form of $r^s_{ph,h}$ introduced in (4.56) and Proposition 4.3.2, we obtain that $C^{s,r}_{ph,h}(\xi)$ could be singular as $\xi \to 0$. We have

$$\lim_{\xi\to 0} r^s_{ph,h}(\xi \pm \pi/h) = \begin{cases} s, & s \in (1,3) \\ 4, & s \in (3,\infty) \end{cases}$$

and $\lim_{\xi\to 0} r^s_{ph,h}(\xi) = 0$, but $\lim_{\xi\to 0} \tan^4(\xi h/2)/r^s_{ph,h}(\xi) = 0$. Then $\lim_{\xi\to 0} C^{s,r}_{ph,h}(\xi) = 0$. We also obtain that

$$\lim_{s\to\infty} C^{s,r}_{ph,h}(\xi) = \Big|\frac{2+\cos(\xi h)}{2-\cos(\xi h)}\Big|^2,$$

which is finite in $\xi \in \Pi_{2h}$. Then there exists $C^{s,r}_{ph,h}$ uniformly as $s \to \infty$. The independence with respect to h follows by rescaling, i.e. from the fact that if $\xi \in \Pi_{2h}$, then $\xi h \in \Pi_2$. As a consequence,

$$C^{s,r}_{ph,h} = \max_{\xi\in\Pi_{2h}} C^{s,r}_{ph,h}(\xi) = \max_{\xi\in\Pi_2} C^{s,r}_{ph,1}(\xi). \tag{4.119}$$

The right hand side in (4.119) does not depend on h. □

Remark 4.5.3. *Before analyzing other filtering mechanisms, let us make some observations concerning the two filtering mechanisms detailed up to this moment.*

1. *According to the microlocal analysis, the time we have obtained in Theorem 4.5.1, $T^{s,\delta}_{ph}$, is not optimal. The optimal one should be*

$$T^{s,\delta}_{ph,*} = \frac{2}{\min\limits_{\xi\in\Pi^\delta_h} \partial_\xi \lambda^s_{ph,h}(\xi)}.$$

   One of the contributions of Theorem 4.5.1 is to find lower and upper bounds of $T^{s,\delta}_{ph}$ according to $T^{s,\delta}_{ph,*}$, for all $\delta \in (0,1)$ and $s > 1$. More precisely, $T^{s,\delta}_{ph,*} \le T^{s,\delta}_{ph} \le 2T^{s,\delta}_{ph,*}$. Then, any improvement of $T^{s,\delta}_{ph}$ in Theorem 4.5.1 improves also the observability time in Theorem 4.5.2.

2. *Both $T^{s,\delta}_{ph,*}$ and $T^{s,\delta}_{ph}$ depend on $\min\limits_{\xi\in\Pi^\delta_h} \partial_\xi \lambda^s_{ph,h}(\xi)$. For $s > 3$, one can guarantee the existence of $\xi_s \in (0,\pi/h)$, s.t. $\partial_\xi \lambda^s_{ph,h}(\xi) > 1$, for $\xi \in (0,\xi_s)$, and $\partial_\xi \lambda^s_{ph,h}(\xi) \in (0,1]$, for $\xi \in [\xi_s,\pi/h)$. Also, $\xi_s \to \pi/h$ as $s \to 3$ and $\xi_s \to \xi_*$, where $\xi_* \in (0,\pi/h)$ is such that $\partial_\xi \lambda_h(\xi_*) = 1$, with λ_h being the dispersion relation corresponding to the P_1-classical finite element method introduced in (4.39). Observe that $\partial_\xi \lambda_h(\pi/2h) = 3\sqrt{3}/4 > 1$. This means $\xi_* > \pi/2h$. Then any Fourier truncation of parameter $\delta \in [1/2, h\xi_s/\pi]$ gives the optimal continuous observability time since $\min\limits_{\xi\in\Pi^\delta_h} \partial_\xi \lambda^s_{ph,h}(\xi) = 1$ and $T^{s,\delta}_{ph,*} = 2$.*

3. *Taking into account that for $s = 3$ the physical group velocity $\partial_\xi \lambda^3_{ph,h}(\xi)$ does not vanish for any wave number $\xi \in \Pi_h$, for all $T > 4$, the observability inequality (4.70) holds uniformly as $h \to 0$ occurs without necessity of any Fourier filtering.*

4. *The same analysis can be carried out on initial data $\widehat{u}^{h,i}$ obtained by a bi-grid algorithm of ratio $n = 1/2^k$, $k \in \mathbb{N}$, i.e. by linear interpolation from a coarse grid of size nh to the fine one.*

4.5.3 Initial data with null jump part + Fourier filtering of the average part

In this subsection and in the following one, the initial data in (4.15), $\overrightarrow{U}^{h,i} = (A^i_j, J^i_j)_{j\in\mathbb{Z}}$, have the following common property concerning the jump part $\overrightarrow{J}^{h,i}$, $i = 0, 1$:

$$\overrightarrow{J}^{h,i} \equiv 0, \ (\text{or } \widehat{J}^{h,i}(\xi) \equiv 0). \tag{4.120}$$

Taking into account (4.33), the total energy corresponding to this choice of the initial data coincides with the total energy corresponding to the classical linear FEM method:

$$E^s_h(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}) = \frac{1}{4\pi} \int\limits_{\Pi_h} \Big[\frac{2 + \cos(\xi h)}{3} |\widehat{A}^{h,1}(\xi)|^2 + \Omega_h(\xi) |\widehat{A}^{h,0}(\xi)|^2 \Big] \, d\xi, \tag{4.121}$$

with Ω_h (being the Fourier symbol of the three-points scheme for the $1-d$ Laplacian) introduced in (4.39).

Then the solution of (4.32) simplifies with respect to the one introduced in (4.47) in the following way:

$$\overrightarrow{U}^h(\xi, t) = \frac{1}{1 - p^s_{ph,h}(\xi) p^s_{sp,h}(\xi)} \begin{pmatrix} 1 & -p^s_{ph,h}(\xi) p^s_{sp,h}(\xi) \\ p^s_{ph,h}(\xi) & -p^s_{ph,h}(\xi) \end{pmatrix} \begin{pmatrix} A^s_{ph,h}(\xi, t) \\ A^s_{sp,h}(\xi, t) \end{pmatrix}, \tag{4.122}$$

where

$$A^s_{ph,h}(\xi, t) = \frac{1}{2} \sum_{\pm} \Big[\widehat{A}^{h,0}(\xi) \pm \frac{\widehat{A}^{h,1}(\xi)}{i\lambda^s_{ph,h}(\xi)} \Big] \exp(\pm it\lambda^s_{ph,h}(\xi)),$$

$$A^s_{sp,h}(\xi, t) = \frac{1}{2} \sum_{\pm} \Big[\widehat{A}^{h,0}(\xi) \pm \frac{\widehat{A}^{h,1}(\xi)}{i\lambda^s_{sp,h}(\xi)} \Big] \exp(\pm it\lambda^s_{sp,h}(\xi))$$

and $p^s_{ph,h}(\xi)$, $p^s_{sp,h}(\xi)$ given by (4.43) and (4.44), respectively.

Remark 4.5.4. *Although the initial data in (4.15), $\overrightarrow{U}^{h,i}$, have null jump part, the property (4.120) is not conserved in time, i.e. for $t > 0$, in general $\overrightarrow{J}^h(t) \neq 0$. However, for all $t > 0$, $\lim\limits_{\xi\to 0} \widehat{J}^h(\xi, t) = 0$. This permits to eliminate the pathological behavior on the spurious diagram at the wave number $\xi = 0$ by finding an uniform upper bound of the total energy of the solutions with respect to the energy concentrated on the physical mode. This, combined with a Fourier truncation on the average part of the initial data $\overrightarrow{A}^{h,i}$ in this subsection or with a bi-grid algorithm of ratio $1/2$ applied to $\overrightarrow{A}^{h,i}$ in the following one, yields two efficient filtering mechanisms allowing to reestablish the semi-discrete uniform observability property.*

The main result of this subsection reads as follows:

Theorem 4.5.3. *Set $\Omega = \{|x| > 1\}$. In (4.15), consider initial data $\overrightarrow{U}^{h,i} = (\overrightarrow{A}^{h,i}, \overrightarrow{J}^{h,i})'$, $i = 0, 1$, having null jump part, i.e. $\overrightarrow{J}^{h,i}$ verifies (4.120), and such that $\overrightarrow{A}^{h,i} \in I^\delta_h$, with $\delta \in (0, 1)$. Then for all $T > T^{s,\delta}_{ph}$, with $T^{s,\delta}_{ph}$ introduced in Theorem 4.5.1, and for all $s > 1$ such that*

$$\max_{\xi\in\Pi^\delta_h} |\lambda^s_{ph,h}(\xi)| < \min_{\xi\in\Pi_h} |\lambda^s_{sp,h}(\xi)|, \tag{4.123}$$

there exists a constant $C_0(T,s,\delta)>0$ independent of h such that the following observability inequality holds:

$$E_h^s(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1}) \leq C_0(T,s,\delta)\int_0^T E_{\Omega,h}^s(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)\,dt. \tag{4.124}$$

Remark 4.5.5. *Let us discuss the range of s within which the condition (4.123) holds for $\delta=1$.*

- *For $s>3$, $\lambda_{ph,h}^s(\xi)$ is strictly increasing and $\lambda_{sp,h}^s(\xi)$ is strictly decreasing on $[0,\pi/h]$, and we have*

$$\min_{\xi\in\Pi_h}|\lambda_{sp,h}^s(\xi)| = \lambda_{sp,h}^s(\pi/h) = 4s > 12 = \max_{\xi\in\Pi_h}|\lambda_{ph,h}^s(\xi)| = \lambda_{ph,h}^s(\pi/h),$$

  meaning that, for $s>3$, condition (4.123) holds.

- *For $s\in(1,3)$, a critical point $\xi_s\in(0,\pi/h)$ could appear on $\lambda_{sp,h}^s(\xi)$, apart from the singular points in $\xi=0,\pi/h$. However, if this holds, ξ_s is a maximum point. Then, the minimum value for $\lambda_{sp,h}^s(\xi)$ is attained at $\xi=0$ or at $\xi=\pi/h$. A simple computation shows that $\lambda_{sp,h}^s(0)=12(s-1)$, $\lambda_{sp,h}^s(\pi/h)=12$ and $\max_{\xi\in\Pi_h}|\lambda_{ph,h}^s(\xi)| = \lambda_{ph,h}^s(\pi/h)=4s$.*

  - For $s\in(2,3)$, $12<12(s-1)$, i.e. the minimum value on $\lambda_{sp,h}^s(\xi)$ is attained at $\xi=\pi/h$ and equals 12. Also, $12>4s$, meaning that the condition (4.123) holds.

  - For $s\in(1,2]$, the minimum value on $\lambda_{sp,h}^s(\xi)$ is attained at $\xi=0$ and equals $12(s-1)$. In order for (4.123) to hold, it is sufficient to impose $12(s-1)>4s$, i.e. $s>3/2$.

  Consequently, (4.123) also holds for $s\in(1.5,3)$.

- *For $s=3$, the condition (4.123) holds with equality and we can not guarantee Theorem 4.5.3 to hold for $\delta=1$, which is the interesting case here taking into account the fact that both group velocities $\partial_\xi\lambda_{ph,h}^s(\xi)$ and $\partial_\xi\lambda_{sp,h}^s(\xi)$ do not vanish at the wave number $\xi=\pi/h$. Step 2 and Step 5 in the proof of Theorem 4.5.3 fail. In fact, the proof fails since $I_h^{1+\epsilon}$ is not defined.*

Before the proof of Theorem 4.5.3, let us introduce some notions and a result that will be used within the proof. If $\overrightarrow{f}^h(t)$ is an evolution process such that its SDFT,

$$\widehat{f}^h(\xi,t) = \widehat{f}_{ph}^h(\xi)\exp(it\lambda_{ph,h}^s(\xi)) + \widehat{f}_{sp}^h(\xi)\exp(it\lambda_{sp,h}^s(\xi)),\ \xi\in\Pi_h, \tag{4.125}$$

involves the two branches of frequencies $\lambda_{ph,h}^s(\xi)$ and $\lambda_{sp,h}^s(\xi)$ and $\widehat{f}_{ph}^h(\xi)$ and $\widehat{f}_{sp}^h(\xi)$ are scalar functions, we define its projection on the physical branch to be

$$\Gamma_{ph}f_j(t) = \frac{1}{2\pi}\int_{\Pi_h}\widehat{f}_{ph}^h(\xi)\exp(it\lambda_{ph,h}^s(\xi))\exp(i\xi x_j)\,d\xi. \tag{4.126}$$

If $\overrightarrow{f}^h(t)$ is a vectorial process, in the sense that both $\widehat{f}_{ph}^h(\xi)$ and $\widehat{f}_{sp}^h(\xi)$ are vectorial functions with the same number of components, we define $\Gamma_{ph}\overrightarrow{f}^h(t)$ to be the vector obtained as Γ_{ph} acting on each one of the components of $\overrightarrow{f}^h(t)$.

On the vectorial process $\overrightarrow{U}^h(t)$ given by (4.122), the projection Γ_{ph} acts as follows

$$\Gamma_{ph}U_j(t) = \frac{1}{2\pi}\int_{\Pi_h}\frac{\sqrt{1+|p_{ph,h}^s(\xi)|^2}}{1-p_{ph,h}^s(\xi)p_{sp,h}^s(\xi)}P_{ph,h}^s(\xi)A_{ph,h}^s(\xi,t)\exp(i\xi x_j)\,d\xi, \tag{4.127}$$

where $P_{ph,h}^s(\xi)$ is the eigenvector corresponding to the physical eigenvalue $\Lambda_{ph,h}^s(\xi)$ introduced in (4.42).

Note that $\Gamma_{ph}\overrightarrow{U}^h(t)$ verifies the wave equation (4.15) with the initial data $\overrightarrow{V}^{h,i}$, $i=0,1$, such that

$$\widehat{V}^{h,i}(\xi)=P^s_{ph,h}(\xi)\frac{\sqrt{1+|p^s_{ph,h}(\xi)|^2}}{1-p^s_{ph,h}(\xi)p^s_{sp,h}(\xi)}\widehat{A}^{h,i}(\xi). \tag{4.128}$$

Then the total energy of $\Gamma_{ph}\overrightarrow{U}^h(t)$, $E^s_h(\Gamma_{ph}\overrightarrow{U}^h(t),\partial_t\Gamma_{ph}\overrightarrow{U}^h(t))$, is conserved in time. We denote it by $E^s_h(\Gamma_{ph}\overrightarrow{U}^{h,0},\Gamma_{ph}\overrightarrow{U}^{h,1})$. Observe that $\widehat{V}^{h,i}(\xi)$, $i=0,1$, verifies the condition (4.53), with

$$\widehat{u}^{h,i}(\xi)=\frac{\sqrt{1+|p^s_{ph,h}(\xi)|^2}}{1-p^s_{ph,h}(\xi)p^s_{sp,h}(\xi)}\widehat{A}^{h,i}(\xi),\ i=0,1.$$

Then the total energy $E^s_h(\Gamma_{ph}\overrightarrow{U}^{h,0},\Gamma_{ph}\overrightarrow{U}^{h,1})$ can be computed using (4.56) and is given by

$$\begin{aligned}&E^s_h(\Gamma_{ph}\overrightarrow{U}^{h,0},\Gamma_{ph}\overrightarrow{U}^{h,1})=\\&\quad=\frac{1}{4\pi}\int\limits_{\Pi_h}\frac{1+|p^s_{ph,h}(\xi)|^2}{|1-p^s_{ph,h}(\xi)p^s_{sp,h}(\xi)|^2}\big(m^s_{ph,h}(\xi)|\widehat{A}^{h,1}(\xi)|^2+r^s_{ph,h}(\xi)|\widehat{A}^{h,0}(\xi)|^2\big)\,d\xi.\end{aligned} \tag{4.129}$$

The following result holds:

Proposition 4.5.2. *In (4.15), consider initial data $\overrightarrow{U}^{h,i}=(\overrightarrow{A}^{h,i},\overrightarrow{J}^{h,i})'$, $i=0,1$, with $\overrightarrow{J}^{h,i}$ satisfying (4.120) and $\overrightarrow{A}^{h,i}\in I^\delta_h$, $\delta\in(0,1)$. Then, for all $s>1$, there exists a constant $C(\delta)>0$ independent of h and s s.t. the following inequality holds*

$$E^s_h(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1})\le C(\delta)E^s_h(\Gamma_{ph}\overrightarrow{U}^{h,0},\Gamma_{ph}\overrightarrow{U}^{h,1}). \tag{4.130}$$

Moreover, for $s\in(1,3)$, $C(\delta)\to\infty$ as $\delta\to1$, and for $s\ge3$, $C(\delta)$ is uniform as $\delta\to1$.

The proof of Proposition 4.5.2 will be provided after the one corresponding to Theorem 4.5.3.

Proof of Theorem 4.5.3. We divide the proof of this theorem in the following steps:

Step 1. Choose $\delta\in(0,1)$ and apply Proposition 4.5.2.

Step 2. Total energy of the projectors P_k when $bc^k<\min\limits_{\xi\in\Pi_h}|\lambda^s_{sp,h}(\xi)|$. For a,b,c,μ as follows

$$1<c<\frac{b-\mu}{a+\mu}<\frac{b}{a}, \tag{4.131}$$

F given by (4.96) and P given by (4.97), we define the projectors P_k by (4.94). Consider $k\in\mathbb{N}$ s.t. for all $\xi\in\Pi^\delta_h$, $bc^k<\min\limits_{\xi\in\Pi^\delta_h}|\lambda^s_{sp,h}(\xi)|$, i.e. $P(\lambda^s_{sp,h}(\xi)/c^k)=0$, for all $\xi\in\Pi^\delta_h$, but for some $\xi\in\Pi^\delta_h$, $\lambda^s_{ph,h}(\xi)\in(ac^k,bc^k)$. Then, taking into account (4.122), we see that

$$P_k\overrightarrow{U}^h(t)=\frac{1}{2\pi}\int\limits_{\Pi^\delta_h}P^s_{ph,h}(\xi)\frac{\sqrt{1+|p^s_{ph,h}(\xi)|^2}}{1-p^s_{ph,h}(\xi)p^s_{sp,h}(\xi)}P\Big(\frac{\lambda^s_{ph,h}(\xi)}{c^k}\Big)A^s_{ph,h}(\xi,t)\exp(i\xi x_j), \tag{4.132}$$

with $P^s_{ph,h}(\xi)$ the eigenvector given by (4.42). This is a solution of the semi-discrete wave equation (4.15) with the initial data $\overrightarrow{W}^{h,i}$, $i=0,1$, such that

$$\widehat{W}^{h,i}(\xi)=P^s_{ph,h}(\xi)\frac{\sqrt{1+|p^s_{ph,h}(\xi)|^2}}{1-p^s_{ph,h}(\xi)p^s_{sp,h}(\xi)}P\Big(\frac{\lambda^s_{ph,h}(\xi)}{c^k}\Big)\widehat{A}^{h,i}(\xi), \tag{4.133}$$

which is of type (4.53), with

$$\widehat{u}^{h,i}(\xi)=\frac{\sqrt{1+|p^s_{ph,h}(\xi)|^2}}{1-p^s_{ph,h}(\xi)p^s_{sp,h}(\xi)}P\Big(\frac{\lambda^s_{ph,h}(\xi)}{c^k}\Big)\widehat{A}^{h,i}(\xi).$$

Then the total energy of the projection $P_k\overrightarrow{U}^h(t)$, $E^s_h(P_k\overrightarrow{U}^h(t),\partial_tP_k\overrightarrow{U}^h(t))$, is conserved in time. Denote this quantity by $E^s_h(P_k\overrightarrow{U}^{h,1},P_k\overrightarrow{U}^{h,1})$. Using (4.56), we get

$$E^s_h(P_k\overrightarrow{U}^{h,1},P_k\overrightarrow{U}^{h,1})=$$
$$\frac{1}{4\pi}\int\limits_{\Pi^\delta_h}\frac{1+|p^s_{ph,h}(\xi)|^2}{|1-p^s_{ph,h}(\xi)p^s_{sp,h}(\xi)|^2}P^2\Big(\frac{\lambda^s_{ph,h}(\xi)}{c^k}\Big)\big(m^s_{ph,h}(\xi)|\widehat{A}^{h,1}(\xi)|^2+r^s_{ph,h}(\xi)|\widehat{A}^{h,0}(\xi)|^2\big)\,d\xi. \quad (4.134)$$

Step 3. Choice the upper bound for the index k of the projectors P_k. As in Step 4 in the proof of Theorem 4.5.2, for $\delta\in(0,1)$ we have fixed, consider k_h such that (4.101) holds. Then, by (4.131), the inequalities (4.102) hold for all $k_0\leq k\leq k_h$. Then any $\lambda^s_{ph,h}(\xi)\in((a+\mu)c^{k_0},\lambda^s_{ph,h}(\delta\pi/h))$ belongs to at least one interval of the form $((a+\mu)c^k,(b-\mu)c^k)$, $k=k_0,\cdots,k_h$, i.e. in the region where $P(\cdots/c^k)\equiv 1$. Therefore, (4.103) holds.

Step 4. Upper bounds of $E^s_h(\Gamma_{ph}\overrightarrow{U}^{h,0},\Gamma_{ph}\overrightarrow{U}^{h,1})$ in terms of the energy of the projectors $E^s_h(P_k\overrightarrow{U}^{h,0},P_k\overrightarrow{U}^{h,1})$. Taking into account the definition of the set $\Pi^{k_0}_h$ provided by (4.104), and of the projection Γ_{k_0} on $\Pi^{k_0}_h$ introduced in (4.105), we obtain the following decomposition of the energy $E^s_h(\Gamma_{ph}\overrightarrow{U}^{h,0},\Gamma_{ph}\overrightarrow{U}^{h,1})$:

$$E^s_h(\Gamma_{ph}\overrightarrow{U}^{h,0},\Gamma_{ph}\overrightarrow{U}^{h,1})=$$
$$=E^s_h(\Gamma_{k_0}\Gamma_{ph}\overrightarrow{U}^{h,0},\Gamma_{k_0}\Gamma_{ph}\overrightarrow{U}^{h,1})+E^s_h((\Gamma^\delta_h-\Gamma_{k_0})\Gamma_{ph}\overrightarrow{U}^{h,0},(\Gamma^\delta_h-\Gamma_{k_0})\Gamma_{ph}\overrightarrow{U}^{h,1}). \quad (4.135)$$

Since $\overrightarrow{A}^{h,i}\in I^\delta_h$, the energies $E^s_h(\Gamma_{ph}\overrightarrow{U}^{h,0},\Gamma_{ph}\overrightarrow{U}^{h,1})$, $E^s_h(\Gamma_{k_0}\Gamma_{ph}\overrightarrow{U}^{h,0},\Gamma_{k_0}\Gamma_{ph}\overrightarrow{U}^{h,1})$ and $E^s_h((\Gamma^\delta_h-\Gamma_{k_0})\Gamma_{ph}\overrightarrow{U}^{h,0},(\Gamma^\delta_h-\Gamma_{k_0})\Gamma_{ph}\overrightarrow{U}^{h,1})$ are given by (4.129), but with Π_h replaced by Π^δ_h, $\Pi^{k_0}_h$ and $\Pi^\delta_h\setminus\Pi^{k_0}_h$, respectively. The first term in (4.135) is a lower order one. For the second one, taking into account (4.103), the explicit for of the energy of the projectors P_k, $k_0\leq k\leq k_h$, (4.134), the condition (4.123) and the fact that $\lambda^s_{ph,h}(\xi)$ is increasing in ξ, we obtain

$$E^s_h((\Gamma^\delta_h-\Gamma_{k_0})\Gamma_{ph}\overrightarrow{U}^{h,0},(\Gamma^\delta_h-\Gamma_{k_0})\Gamma_{ph}\overrightarrow{U}^{h,1})\leq\sum_{k=k_0}^{k_h}E^s_h(P_k\overrightarrow{U}^{h,0},P_k\overrightarrow{U}^{h,1}). \quad (4.136)$$

From (4.135) and (4.136), we get

$$E^s_h(\Gamma_{ph}\overrightarrow{U}^{h,0},\Gamma_{ph}\overrightarrow{U}^{h,1})\leq E^s_h(\Gamma_{k_0}\Gamma_{ph}\overrightarrow{U}^{h,0},\Gamma_{k_0}\Gamma_{ph}\overrightarrow{U}^{h,1})+\sum_{k=k_0}^{k_h}E^s_h(P_k\overrightarrow{U}^{h,0},P_k\overrightarrow{U}^{h,1}). \quad (4.137)$$

Step 5. $P_k\overrightarrow{U}^{h,i}\in(I^{\delta+\epsilon}_h)^2$, $i=0,1$. This step is for free, without necessity of the arguments in Step 6 within the proof of Theorem 4.5.2, due to the fact that $\overrightarrow{A}^{h,i}\in I^\delta_h\subset I^{\delta+\epsilon}_h$ and then $P_k\overrightarrow{U}^{h,i}\in(I^{\delta+\epsilon}_h)^2$.

Step 6. Observability inequality for the projections $P_k\overrightarrow{U}^h(t)$. Using the hypothesis on T in the statement of the theorem and Step 1 in the proof of Theorem 4.5.2 and more precisely (4.93), in which we are allowed to introduce $P_k\overrightarrow{U}^{h,i}\in(I^{\delta+\epsilon}_h)^2$, we obtain (4.110), for all $k_0\leq k\leq k_h$.

Step 7. Observability inequality for $\overrightarrow{U}^h(t)$. This reduces to apply Lemma 4.5.2 in the same setting we have established in the proof of Theorem 4.5.2, in Step 8.

Step 8. Choice of k_0. Taking into account Proposition 4.5.2, i.e. (4.130), the inequalities (4.137), (4.110), Lemma 4.5.2 and the direct inequality (4.112), we choose k_0 such that

$$\frac{C(\delta)C_{ph}^{s,\delta+\epsilon}(T-4\gamma)C(P,T,\gamma)T}{c^{2k_0}} \leq \frac{1}{2}, \tag{4.138}$$

with $C(\delta)$, $C_{ph}^{s,\delta}(T)$, $C(P,T,\gamma)$ introduced in Proposition 4.5.2, Theorem 4.5.1 and Lemma 4.5.2, respectively, all being constants independent of h. Consequently,

$$\begin{aligned} &E_h^s(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1}) \leq \\ &\leq \widetilde{C}_0(T,s,\delta,\epsilon,\gamma,P,c)\int_0^T E_{\Omega,h}^s(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)\,dt + 2C(\delta)E_h^s(\Gamma_{k_0}\Gamma_{ph}\overrightarrow{U}^{h,0},\Gamma_{k_0}\Gamma_{ph}\overrightarrow{U}^{h,1}), \end{aligned} \tag{4.139}$$

where $\widetilde{C}_0(T,s,\delta,\epsilon,\gamma,P,c) = 2C(\delta)C_{ph}^{s,\delta+\epsilon}(T-4\gamma)C(P,c)$, with $C(P,c)$ the constant independent of h, introduced in Lemma 4.5.2.

Step 9. Compactness of the term $E_h^s(\Gamma_{k_0}\Gamma_{ph}\overrightarrow{U}^{h,0},\Gamma_{k_0}\Gamma_{ph}\overrightarrow{U}^{h,1})$.

This concludes the proof of Theorem 4.5.3. □

Proof of Proposition 4.5.2. Taking into account (4.121) and (4.129), in which Π_h has to be replaced by Π_h^δ because $\overrightarrow{A}^{h,i} \in I_h^\delta$, we see that (4.130) is equivalent to the existence of:

$$\|Cm_h^i\|_{L^\infty_{s,\xi}((1,\infty)\times\Pi_h^\delta)} \text{ and } \|Cm_h^i\|_{L^\infty_{s,\xi}((1,\infty)\times\Pi_h^\delta)} \tag{4.140}$$

where

$$Cm_h^s(\xi) = \frac{(2+\cos(\xi h))|1-p_{ph,h}^s(\xi)p_{sp,h}^s(\xi)|^2}{3m_{ph,h}^s(\xi)(1+|p_{ph,h}^s(\xi)|^2)}$$

and

$$Cr_h^s(\xi) = \frac{\Omega_h(\xi)|1-p_{ph,h}^s(\xi)p_{sp,h}^s(\xi)|^2}{r_{ph,h}^s(\xi)(1+|p_{ph,h}^s(\xi)|^2)}.$$

and to the fact that these quantities do not depend on h.

Taking into account the explicit form of $m_{ph,h}^s(\xi)$, $r_{ph,h}^s(\xi)$ introduced in (4.56), $Cm_h^s(\xi)$ and $Cr_h^s(\xi)$ simplify as follows

$$Cm_h^s(\xi) = \frac{(2+\cos(\xi h))|1-p_{ph,h}^s(\xi)p_{sp,h}^s(\xi)|^2}{1+\frac{1}{4}|p_{ph,h}^s(\xi)|^2+2\Big|\cos\Big(\frac{\xi h}{2}\Big)+\frac{i}{2}\sin\Big(\frac{\xi h}{2}\Big)p_{ph,h}^s(\xi)\Big|^2} \text{ and } Cr_h^s(\xi) = 1-p_{ph,h}^s(\xi)p_{sp,h}^s(\xi).$$

Observe that $Cm_h^s(\xi)$ and $Cr_h^s(\xi)$ are rational expressions depending on trigonometric functions and on s, so they are continuous in s and ξ, excepting eventually the possible singularities. Before studying these possible singularities of $Cm_h^s(\xi)$ and $Cr_h^s(\xi)$, observe that if $\|Cm_h^s\|_{L^\infty(\Pi_h^\delta)}$ and $\|Cr_h^s\|_{L^\infty(\Pi_h^\delta)}$ exists, then, by a rescaling argument,

$$\|Cm_h^s\|_{L^\infty(\Pi_h^\delta)} = \|Cm_1^s\|_{L^\infty(\Pi_1^\delta)} \text{ and } \|Cr_h^s\|_{L^\infty(\Pi_h^\delta)} = \|Cr_1^s\|_{L^\infty(\Pi_1^\delta)}$$

so they do not depend on h.

Possible singularities of $Cm_h^s(\xi)$ in ξ, with fixed $s>1$. The denominator in $Cm_1^s(\xi)$ is strictly positive. We have to analyze the possibility that the numerator is infinite. The numerator is infinite iff $-p_{ph,1}^s(\xi)p_{sp,1}^s(\xi)$ is infinite. Using the explicit expressions of $p_{ph,1}^s(\xi)$ and $p_{sp,1}^s(\xi)$ provided by (4.43) and (4.44), we get

$$\begin{aligned} &-p_{ph,1}^s(\xi)p_{sp,1}^s(\xi) = \\ &= \frac{1}{4\cos^2\Big(\frac{\xi}{2}\Big)\Big(s-\cos^2\Big(\frac{\xi}{2}\Big)\Big)}\Big[2-\cos(\xi)-\frac{12\Big(s-\cos^2\Big(\frac{\xi}{2}\Big)\Big)}{12+2(s-3)(2+\cos(\xi))+\sqrt{\Delta_1^s(\xi)}}\Big]^2. \end{aligned} \tag{4.141}$$

Taking into account that $s>1$, the product $-p_{ph}^s(\xi)p_{sp}^s(\xi)$ could be infinity iff $\xi\to\pm\pi$. By an analysis similar to the one done in the proof of Propositions 4.3.2 or 4.3.3, we see that

$$-\lim_{\xi\to\pm\pi} p_{ph}^s(\xi)p_{sp}^s(\xi)=\begin{cases}+\infty, & s\in(1,3)\\ 1, & s=3\\ 0, & s>3.\end{cases} \tag{4.142}$$

By the property (Sp4) of Proposition 4.3.3, we see that for $s\in(1,3)$,

$$\lim_{\xi\to\pm\pi} Cm_1^s(\xi)=+\infty.$$

However, when $\xi\in\Pi_1^\delta$, $\delta\in(0,1)$, and $s\in(1,3)$, $Cm_1^s(\xi)$ has no singularities. Then, for any $\delta\in(0,1)$, $Cm_1^s\in L^\infty(\Pi_1^\delta)$, with

$$\|Cm_1^s\|_{L^\infty(\Pi_1^\delta)}\to\begin{cases}+\infty, & s\in(1,3)\\ \text{finite limit}, & s\geq 3\end{cases}$$

as $\delta\to 1$.

Possible singularities of $Cm_1^s(\xi)$ in s, with fixed $\xi\in\Pi_1$. Taking into account that $s>1$, $Cm_1^s(\xi)$ could be singular in s only when for some value of s $-p_{ph,1}^s(\xi)p_{sp,1}^s(\xi)$ is infinite. This could happen only as $s\to\infty$. Nevertheless, $\lim_{s\to\infty} Cm_1^s(\xi)=1$. This analysis shows that

$$\|Cm_1^s\|_{L^\infty(\Pi_1^\delta)}\leq\begin{cases}C_m^1(\delta)=\|Cm_1^{\cdot}\|_{L^\infty_{s,\xi}((1,3)\times\Pi_1^\delta)}, & s\in(1,3)\\ C_m^2=\|Cm_1^{\cdot}\|_{L^\infty_{s,\xi}((3,\infty)\times\Pi_1)}, & s\geq 3.\end{cases} \tag{4.143}$$

The analysis of the singularities of $Cr_1^s(\xi)$ is even simpler than the one of $Cm_1^s(\xi)$, based in principal on (4.142). We conclude that

$$\|Cr_1^s\|_{L^\infty(\Pi_1^\delta)}\leq\begin{cases}C_r^1(\delta)=\|Cr_1^{\cdot}\|_{L^\infty_{s,\xi}((1,3)\times\Pi_1^\delta)}, & s\in(1,3)\\ C_r^2=\|Cr_1^{\cdot}\|_{L^\infty_{s,\xi}((3,\infty)\times\Pi_1)}, & s\geq 3.\end{cases} \tag{4.144}$$

Consequently, (4.130) holds with

$$C(\delta)=\begin{cases}\max\{C_m^1(\delta),C_r^1(\delta)\}, & s\in(1,3)\\ \max\{C_m^2,C_r^2\}, & s\geq 3.\end{cases}$$

□

4.5.4 Initial data with null jump part + bi-grid algorithm on the average part

In this subsection, we deal with initial data in (4.15), $\overrightarrow{U}^{h,i}=(\overrightarrow{A}^{h,i},\overrightarrow{J}^{h,i})'$, $i=0,1$, which apart from the condition (4.120), satisfy the fact that the average part $\overrightarrow{A}^{h,i}$ is obtained by a bi-grid algorithm, i.e. by linear interpolation from a coarse grid, G^{2h}, to a finer one, G^h. More precisely, the sequences $\overrightarrow{A}^{h,i}$ verify the condition

$$A_{2j}^i=\frac{A_{2j+1}^i+A_{2j-1}^i}{2},\ \forall j\in\mathbb{Z},\ \forall i=0,1. \tag{4.145}$$

The main result of this subsection is as follows (cf. [41]):

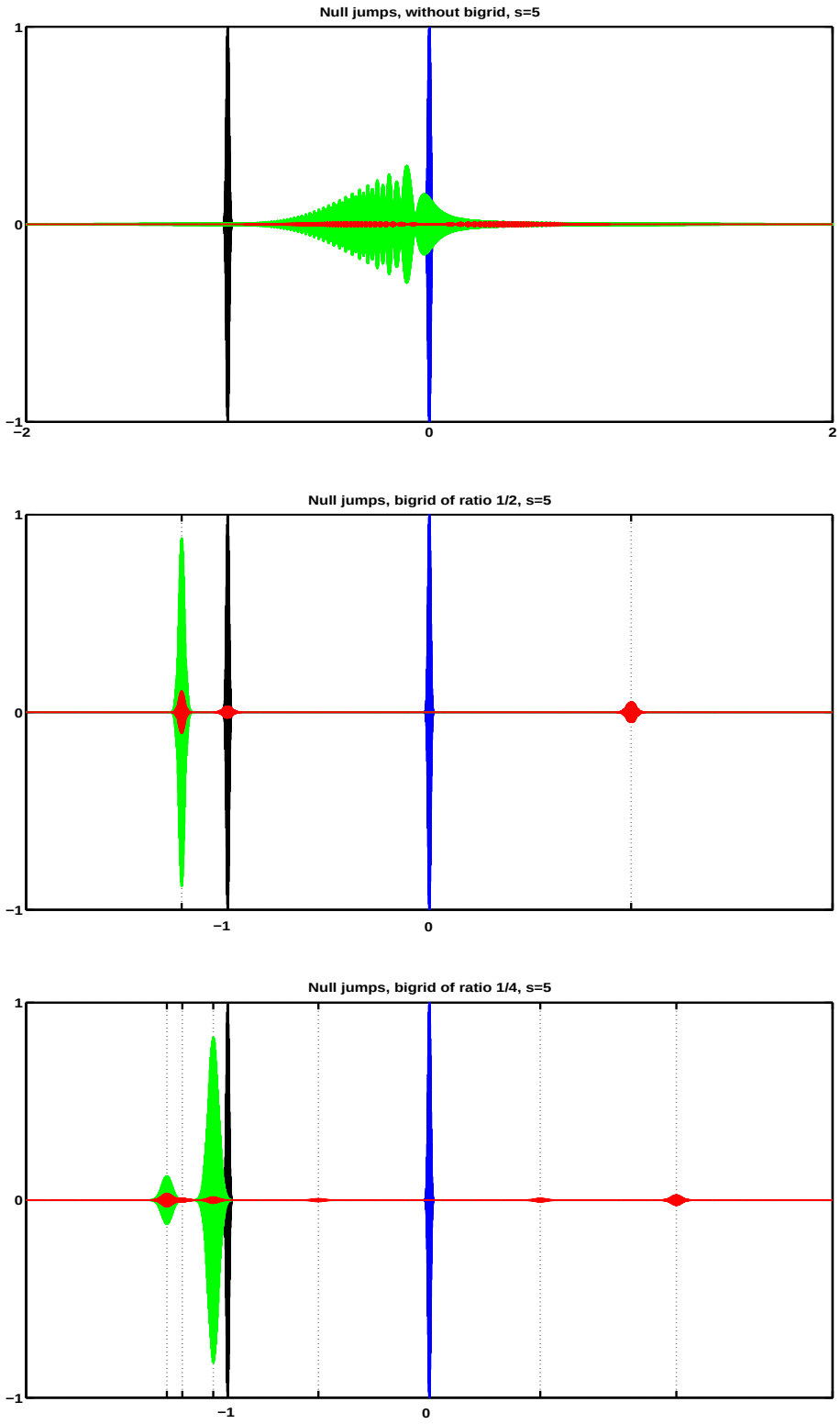

Figure 4.9: Propagation of waves packets for the SIPG semi-discrete wave equations with different types of bi-grids.

Theorem 4.5.4. *Set $\Omega = \{x : |x| > 1\}$. In (4.15) consider initial data $\overrightarrow{U}^{h,i} = (\overrightarrow{A}^{h,i}, \overrightarrow{J}^{h,i})'$, $i = 0, 1$, satisfying the conditions (4.120) and (4.145). For all $T > T_{ph}^{s,1/2}$, with $T_{ph}^{s,\delta}$ given in Theorem 4.5.1, and all $s > 1$ verifying (4.123) with $\delta = 1/2$, there exists a constant $C_{0,bigrid}(T,s) > 0$ independent of h such that the following observability inequality holds:*

$$E_h^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}) \leq C_{0,bigrid}^s(T,s) \int_0^T E_{\Omega,h}^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\, dt. \tag{4.146}$$

In Figure 4.9, we represent the two components of the solution of the SIPG semi-discretization of the wave equation: $\overrightarrow{A}(t)$ (green) and $\overrightarrow{J}(t)$ (red) for $s = 5$, at time $t = 1$, starting with $\widehat{A}^{h,0}(\xi) = \sqrt{2\pi/\gamma^2}\exp(-(\xi - \pi/h)^2/(2\gamma^2))$, $\widehat{A}^{h,1}(\xi) = i\lambda_{ph,h}(\xi)\widehat{A}^{h,0}(\xi)$, $\widehat{J}^{h,0}(\xi) = \widehat{J}^{h,1}(\xi) = 0$, $h = 0.001$ and $\gamma = h^{-2/3}$, without bigrid (top), with bigrid of ratio 1/2 (middle) and bigrid of ratio 1/4 (bottom).

In the proof of Theorem 4.5.4, we will use the following result:

Proposition 4.5.3. *In (4.15), consider an initial data $\overrightarrow{U}^{h,i} = (\overrightarrow{A}^{h,i}, \overrightarrow{J}^{h,i})'$, $i = 0, 1$, satisfying both conditions (4.120) and (4.145). Then, for all $s > 1$, the following estimate holds*

$$E_h^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}) \leq 2E_h^s(\Gamma_h^{1/2}\overrightarrow{U}^{h,0}, \Gamma_h^{1/2}\overrightarrow{U}^{h,1}), \tag{4.147}$$

where Γ_h^δ is the projection defined by (4.68).

Proof of Proposition 4.5.3. From the fact that the average part of the initial data $\overrightarrow{A}^{h,i}$ verifies (4.145), by Lemma 4.5.3, we have

$$\widehat{A}^{h,i}(\xi) = \cos^2\Big(\frac{\xi h}{2}\Big)\widehat{A}^{2h,i}(\xi), \forall \xi \in \Pi_h, i = 0, 1, \tag{4.148}$$

where $\widehat{A}^{2h,i}(\xi)$ is the SDFT at the scale $2h$ of the sequence $(A_{2j+1}^i)_{j\in\mathbb{Z}}$, which is a π/h-periodic function. Taking into account the fact that the initial data $\overrightarrow{U}^{h,i}$, $i = 0, 1$, verify (4.120), the total energy $E_h^s(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1})$ is given by (4.121). Also (4.118) holds. Taking into account the π/h-periodicity of $\widehat{A}^{h,i}$, $i = 0, 1$, we get

$$\begin{aligned} &E_h^s((\Gamma_h^1 - \Gamma_h^{1/2})\overrightarrow{U}^{h,0}, (\Gamma_h^1 - \Gamma_h^{1/2})\overrightarrow{U}^{h,1}) = \\ &\quad = \frac{1}{4\pi}\int_{\Pi_{2h}} \Big[\widetilde{m}_h(\xi)\frac{2 + \cos(\xi h)}{3}|\widehat{A}^{h,1}(\xi)|^2 + \widetilde{r}_h(\xi)\Omega_h(\xi)|\widehat{A}^{h,0}(\xi)|^2\Big]\, d\xi, \end{aligned} \tag{4.149}$$

where

$$\widetilde{m}_h(\xi) = \frac{2 - \cos(\xi h)}{2 + \cos(\xi h)}\tan^4\Big(\frac{\xi h}{4}\Big) \text{ and } \widetilde{r}_h(\xi) = \tan^2\Big(\frac{\xi h}{2}\Big).$$

An easy computation show that both $\widetilde{m}_h$, $\widetilde{r}_h$ are increasing functions on Π_{2h}, therefore their maximum value is attained in $\pi/2h$ and $\|\widetilde{m}_h\|_{L^\infty(\Pi_{2h})} = \|\widetilde{r}_h\|_{L^\infty(\Pi_{2h})} = 1$. This means that

$$E_h^s((\Gamma_h^1 - \Gamma_h^{1/2})\overrightarrow{U}^{h,0}, (\Gamma_h^1 - \Gamma_h^{1/2})\overrightarrow{U}^{h,1}) \leq E_h^s(\Gamma_h^{1/2}\overrightarrow{U}^{h,0}, \Gamma_h^{1/2}\overrightarrow{U}^{h,1}).$$

The conclusion of Proposition 4.5.3 follows using this last estimate in (4.118). □

In the following, we sketch the proof of the main result of this subsection, Theorem 4.5.4.

Proof of Theorem 4.5.4. We divide this proof into several steps as follows.

Step 1. Apply Proposition 4.5.3, so the estimate (4.147) holds.

Step 2. Apply Proposition 4.5.2 with $\delta = 1/2$ and initial data $\Gamma_h^{1/2}\overrightarrow{U}^{h,i}$, $i = 0, 1$. Then

$$E_h^s(\Gamma_h^{1/2}\overrightarrow{U}^{h,0}, \Gamma_h^{1/2}\overrightarrow{U}^{h,1}) \leq C(1/2)E_h^s(\Gamma_{ph}\Gamma_h^{1/2}\overrightarrow{U}^{h,0}, \Gamma_{ph}\Gamma_h^{1/2}\overrightarrow{U}^{h,1}), \tag{4.150}$$

with $C(1/2) > 0$ being the constant independent of h introduced in Proposition 4.5.2.

Step 3. Total energy of the projectors P_k when $bc^k \leq \min_{\xi\in\Pi_h} |\lambda_{ph,h}^s(\xi)|$. See Step 2 in the proof of Theorem 4.5.3 with $\delta = 1/2$.

Step 4. Choice of the upper bound for the index k of the projectors. See Step 4 in the proof of Theorem 4.5.2, with $\delta = 1/2$.

Step 5. Upper bound of $E_h^s(\Gamma_{ph}\Gamma_h^{1/2}\overrightarrow{U}^{h,0}, \Gamma_{ph}\Gamma_h^{1/2}\overrightarrow{U}^{h,1})$ in terms of the energy of the projectors $E_h^s(P_k\overrightarrow{U}^{h,0}, P_k\overrightarrow{U}^{h,0})$. Taking into account the definition of the set $\Pi_h^{k_0}$, (4.104), and of the projection Γ_{k_0} on it, (4.105), we have the following decomposition of the energy $E_h^s(\Gamma_{ph}\Gamma_h^{1/2}\overrightarrow{U}^{h,0}, \Gamma_{ph}\Gamma_h^{1/2}\overrightarrow{U}^{h,1})$:

$$\begin{aligned} &E_h^s(\Gamma_{ph}\Gamma_h^{1/2}\overrightarrow{U}^{h,0}, \Gamma_{ph}\Gamma_h^{1/2}\overrightarrow{U}^{h,1}) = \\ &= E_h^s(\Gamma_{ph}\Gamma_{k_0}\overrightarrow{U}^{h,0}, \Gamma_{ph}\Gamma_{k_0}\overrightarrow{U}^{h,1}) + E_h^s(\Gamma_{ph}(\Gamma_h^{1/2} - \Gamma_{k_0})\overrightarrow{U}^{h,0}, \Gamma_{ph}(\Gamma_h^{1/2} - \Gamma_{k_0})\overrightarrow{U}^{h,1}). \end{aligned} \tag{4.151}$$

By the same arguments as in Step 4 in the proof of Theorem 4.5.3, the estimate (4.136) holds with $\delta = 1/2$. From this and (4.151), we get

$$E_h^s(\Gamma_{ph}\Gamma_h^{1/2}\overrightarrow{U}^{h,0}, \Gamma_{ph}\Gamma_h^{1/2}\overrightarrow{U}^{h,1}) \leq E_h^s(\Gamma_{ph}\Gamma_{k_0}\overrightarrow{U}^{h,0}, \Gamma_{ph}\Gamma_{k_0}\overrightarrow{U}^{h,1}) + \sum_{k=k_0}^{k_h} E_h^s(P_k\overrightarrow{U}^{h,0}, P_k\overrightarrow{U}^{h,1}). \tag{4.152}$$

Step 6. $P_k\overrightarrow{U}^{h,0} \in (I_h^{1/2+\epsilon})^2$. See Step 6 in the proof of Theorem 4.5.2.

Step 7. Observability inequality for the projections $P_k\overrightarrow{U}^h(t)$. See Step 7 in the proof of Theorem 4.5.2 and inequality (4.93) with $\delta = 1/2$.

Step 8. Observability inequality for the solution $\overrightarrow{U}^h(t)$. Apply Lemma 4.5.2 in the setting we established in Step 8 in the proof of Theorem 4.5.2.

Step 9. Choice of k_0. Taking into account (4.147), (4.150), (4.152), (4.93) with $\delta = 1/2$, Lemma 4.5.2 and the direct inequality (4.112), we choose k_0 s.t. (4.138) holds, with $\delta = 1/2$ and $1/4$ in the right hand side. Then (4.139) holds with $\delta = 1/2$ and the right hand side multiplied by 2.

Step 10. Compactness of the term $E_h^s(\Gamma_{ph}\Gamma_{k_0}\overrightarrow{U}^{h,0}, \Gamma_{ph}\Gamma_{k_0}\overrightarrow{U}^{h,1})$. □

4.6 Fourier analysis of the LDG methods

Before proceeding with the Fourier analysis of other symmetric DG methods, let us give a preliminary result concerning the inverse of the mass matrix M_h:

Lemma 4.6.1. *The inverse of the mass matrix M_h given by (4.16) corresponding to a P_1-DG method is a block-tri-diagonal matrix generated by the stencil:*

$$m_h^{-1} = \frac{1}{h}\left(\begin{array}{cc|cc|cc} -\frac{1}{2} & 1 & 2 & 0 & -\frac{1}{2} & -1 \\ -1 & 2 & 0 & 8 & 1 & 2 \end{array}\right) \tag{4.153}$$

Proof. In order to find the inverse of M_h, for an arbitrary sequence $\overrightarrow{g} = (g_j^A, g_j^J)_{j\in\mathbb{Z}}$, we have to solve the system $M_h\overrightarrow{f}' = \overrightarrow{g}'$, with $\overrightarrow{f} = (f_j^A, f_j^J)_{j\in\mathbb{Z}}$. Explicitly, $\overrightarrow{f}$ is the solution of the following infinite coupled system of linear equations:

$$\left\{\begin{array}{ll} \frac{h}{6}f_{j-1}^A - \frac{h}{12}f_{j-1}^J + \frac{2h}{3}f_j^A + \frac{h}{6}f_{j+1}^A + \frac{h}{12}f_{j+1}^J = g_j^A & (I) \\ \frac{h}{12}f_{j-1}^A - \frac{h}{24}f_{j-1}^J + \frac{h}{6}f_j^J - \frac{h}{12}f_{j+1}^A - \frac{h}{24}f_{j+1}^J = g_j^J & (II). \end{array}\right. \tag{4.154}$$

The system (4.154, I.-II.) is equivalent to the following one, obtained as (I+2II) and (I-2II):

$$\left\{\begin{array}{l} \frac{h}{3}f_{j-1}^A - \frac{h}{6}f_{j-1}^J + \frac{2h}{3}f_j^A + \frac{h}{3}f_j^J = g_j^A + 2g_j^J \quad (III) \\ \frac{2h}{3}f_j^A - \frac{h}{3}f_j^J + \frac{h}{3}f_{j+1}^A + \frac{h}{6}f_{j+1}^J = g_j^A - 2g_j^J \quad (IV). \end{array}\right. \tag{4.155}$$

By shifting the equation (4.155-II.) by one unit to the left, we have the following system in the unknowns $2f_{j-1}^A - f_{j-1}^J$ and $2f_j^A + f_j^J$

$$\left\{\begin{array}{l} \frac{h}{6}\big(2f_{j-1}^A - f_{j-1}^J\big) + \frac{h}{3}\big(2f_j^A + f_j^J\big) = g_j^A + 2g_j^J \\ \frac{h}{3}\big(2f_{j-1}^A - f_{j-1}^J\big) + \frac{h}{6}\big(2f_j^A + f_j^J\big) = g_{j-1}^A - 2g_{j-1}^J, \end{array}\right. \tag{4.156}$$

whose solution is

$$\left\{\begin{array}{l} \frac{h}{2}\big(2f_{j-1}^A - f_{j-1}^J\big) = 2\big(g_{j-1}^A - 2g_{j-1}^J\big) - \big(g_j^A + 2g_j^J\big) \quad (V) \\ \frac{h}{2}\big(2f_j^A + f_j^J\big) = -\big(g_{j-1}^A - 2g_{j-1}^J\big) + 2\big(g_j^A + 2g_j^J\big) \quad (VI). \end{array}\right. \tag{4.157}$$

By shifting the first equation to the right by one unit, (4.157) becomes

$$\left\{\begin{array}{l} \frac{h}{2}\big(2f_j^A - f_j^J\big) = 2\big(g_j^A - 2g_j^J\big) - \big(g_{j+1}^A + 2g_{j+1}^J\big) \\ \frac{h}{2}\big(2f_j^A + f_j^J\big) = -\big(g_{j-1}^A - 2g_{j-1}^J\big) + 2\big(g_j^A + 2g_j^J\big), \end{array}\right. \tag{4.158}$$

whose solution is given by

$$\left\{\begin{array}{l} 2hf_j^A = -\big(g_{j-1}^A - 2g_{j-1}^J\big) + 4g_j^A - \big(g_{j+1}^A + 2g_{j+1}^J\big) \\ hf_j^J = -\big(g_{j-1}^A - 2g_{j-1}^J\big) + 8g_j^J + \big(g_{j+1}^A + 2g_{j+1}^J\big), \end{array}\right.$$

which concludes the proof of the form of (4.153). □

Let us introduce the following bilinear forms defined on $V_h \times V_h$:

$$\begin{aligned} e_h(u,v) = \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}} r_{x_j}([u])(x) r_{x_j}([v])(x)\,dx, \quad f_h^\beta(u,v) = -\beta \sum_{j\in\mathbb{Z}} ([u](x_j)[v_x](x_j) + [v](x_j)[u_x](x_j)) \\ + \int_{\mathbb{R}} (r([u])(x) + \beta l([u])(x))(r([v])(x) + \beta l([v])(x))\,dx, \end{aligned} \tag{4.159}$$

where r_{x_j}, r and l are the lift operators defined as follows:

$$\int_{\mathbb{R}} r_{x_j}(u)(x)v(x)\,dx = -u(x_j)\{v\}(x_j) \tag{4.160}$$

and

$$\int_{\mathbb{R}} r(u)(x)v(x)\,dx = -\sum_{j\in\mathbb{Z}} u(x_j)\{v\}(x_j), \quad \int_{\mathbb{R}} l(u)(x)v(x)\,dx = -\sum_{j\in\mathbb{Z}} u(x_j)[v](x_j). \tag{4.161}$$

All the examples of symmetric bilinear form associated to the DG methods for elliptic problems presented in [6] can be written in the unified form:

$$a_h^{s,\mu,\alpha,\beta}(u,v) = b_h(u,v) + c_h(u,v) + \frac{s}{h}d_h(u,v) + \mu e_h(u,v) + \alpha f_h^\beta(u,v), \tag{4.162}$$

with b_h, c_h, d_h defined by (4.11) and $\alpha \in \{0,1\}$. We identify the following particular cases:

- $(\mu = 0, \alpha = 0, s > 1)$ corresponds to the SIPG method introduced in [5].
- $(s = 0, \mu > 1/2, \alpha = 0)$ corresponds to Bassi and al. [9]. In fact, we will see that in dimension $d = 1$, the methods SIPG and Bassi and al. are completely equivalent.

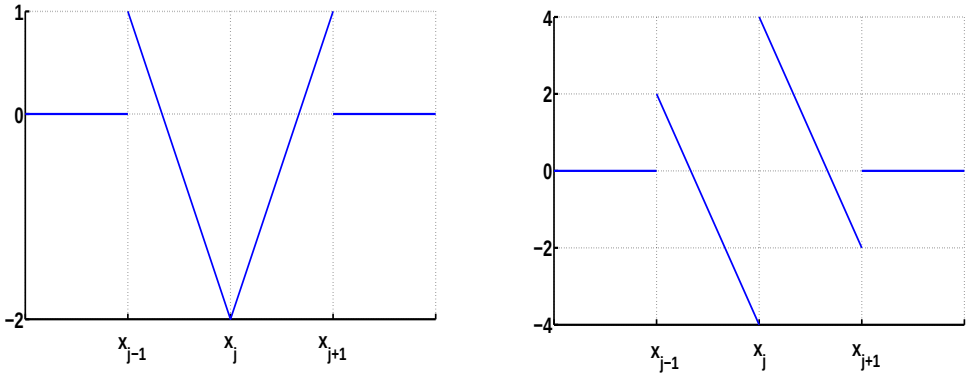

Figure 4.10: The functions ϕ_j (left) and $\widetilde{\phi}_j$ (right).

- $(s = 0, \mu = 0, \alpha = 1, \beta = 0)$ corresponds to the Bassi and Rebay method, introduced in [8], which is not coercive.
- $(s > 0, \mu = 0, \alpha = 1, \beta \in \mathbb{R})$ corresponds to the LDG method introduced in [16] and [22].

The following result characterizes the lift operator $r_{x_j}([\cdot])$ defined by (4.160), for all $j \in \mathbb{Z}$.

Proposition 4.6.1. *For all $j \in \mathbb{Z}$ and all $u \in V_h$ admitting the decomposition $u(x) = \sum_{k\in\mathbb{Z}}(u_k^A \phi_k^A(x) + u_k^J \phi_k^J(x))$, the lift operator $r_{x_j}([u]) \in V_h$ has the form*

$$r_{x_j}([u])(x) = \frac{u_j^J}{h}\phi_j(x), \tag{4.163}$$

where

$$\begin{aligned}\phi_j(x) &= \frac{1}{2}\phi_{j-1}^A(x) - \phi_{j-1}^J(x) - 2\phi_j^A(x) + \frac{1}{2}\phi_{j+1}^A(x) + \phi_{j+1}^J(x)\\ &= \begin{cases} -\frac{x-x_j}{h} - 2\frac{x-x_{j-1}}{h}, & x \in (x_{j-1}, x_j)\\ 2\frac{x-x_{j+1}}{h} + \frac{x-x_j}{h}, & x \in (x_j, x_{j+1})\\ 0, & \textit{otherwise.}\end{cases}\end{aligned} \tag{4.164}$$

Proof. Let us fix $j \in \mathbb{Z}$. Since $r_{x_j}([u]) \in V_h$, we can write it as $r_{x_j}([u])(x) = \sum_{k\in\mathbb{Z}}(a_k^A \phi_k^A(x) + a_k^J \phi_k^J(x))$, with $\overrightarrow{a} = (a_k^A, a_k^J)_{k\in\mathbb{Z}}$ to be determined in what follows. On the other hand, by considering ϕ_l^A and ϕ_l^J, $l \in \mathbb{Z}$, as test functions in (4.160), we obtain $M_h \overrightarrow{a}' = \overrightarrow{f}'$, where $\overrightarrow{f} = (f_j^A, f_l^J)_{l\in\mathbb{Z}} = (-u_j^J \delta_{j,l}, 0)_{l\in\mathbb{Z}}$. Clearly, $\overrightarrow{a}$ depends on j, but to make the things clearer, we do not emphasize this fact in the notation. By making use of the explicit form of the inverse of the mass matrix M_h given by the stencil 4.153, we obtain that

$$\begin{pmatrix} a_l^A \\ a_l^J \end{pmatrix} = \frac{u_j^J}{h}\begin{pmatrix} \frac{1}{2}\delta_{j,l-1} - 2\delta_{j,l} + \frac{1}{2}\delta_{j,l+1} \\ \delta_{j,l-1} - \delta_{j,l+1} \end{pmatrix}.$$

In this way, $(a_{j-1}^A, a_{j-1}^J) = u_j^J/h(1/2, -1)$, $(a_j^A, a_j^J) = u_j^J/h(-2, 0)$ and $(a_{j+1}^A, a_{j+1}^J) = u_j^J/h(1/2, 1)$. Otherwise, for all $l \in \mathbb{Z}$ with $|l - j| \geq 2$, we have $(a_l^A, a_l^J) = (0, 0)$. Then the conclusion of Proposition 4.6.1 follows. □

This means that the stencil generated by the bilinear form e_h is

$$r_{e_h} = \left(\begin{array}{cc|cc|cc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2}{h} & 0 & 0 \end{array}\right) = \frac{2}{h} r_{d_h}.$$

Given the relation between the stencils r_{e_h} and r_{d_h} and denoting by

$$\widetilde{a}^{s,\alpha,\beta}(u,v)=b_h(u,v)+c_h(u,v)+\frac{s}{h}d_h(u,v)+\alpha f_h^{\beta}(u,v)=a_h^s(u,v)+\alpha f_h^{\beta}(u,v),$$

with a_h^s the bilinear form associated to the SIPG method introduced in (4.13), we obtain that

$$a_h^{s,\mu,\alpha,\beta}(u,v)=\widetilde{a}^{s+2\mu,\alpha,\beta}(u,v),\forall u,v\in V_h.$$

In what follows, we give the characterization of the lift operators $r([\cdot])$ and $l([\cdot])$.

Proposition 4.6.2. *For all $u\in V_h$ having the decomposition $u(x)=\sum\limits_{k\in\mathbb{Z}}(u_k^A\phi_k^A(x)+u_k^J\phi_k^J(x))$ the lift operators $r([u]),l([u])\in V_h$ have the form:*

$$r([u])(x)=\frac{1}{h}\sum_{k\in\mathbb{Z}}u_k^J\phi_k(x)=\frac{1}{h}\sum_{k\in\mathbb{Z}}\Big[\Big(\frac{1}{2}u_{k-1}^J-2u_k^J+\frac{1}{2}u_{k+1}^J\Big)\phi_k^A(x)+(u_{k-1}^J-u_{k+1}^J)\phi_k^J(x)\Big] \quad (4.165)$$

and

$$l([u])(x)=\frac{1}{h}\sum_{k\in\mathbb{Z}}u_k^J\widetilde{\phi}_k(x)=\frac{1}{h}\sum_{k\in\mathbb{Z}}\big[(u_{k+1}^J-u_{k-1}^J)\phi_k^A(x)-(2u_{k+1}^J+8u_k^J+2u_{k-1}^J)\phi_k^J(x)\big], \quad (4.166)$$

where $\phi_k(x)$ is defined by (4.164) and $\widetilde{\phi}_k(x)$ is defined by

$$\begin{aligned}\widetilde{\phi}_k(x)&=\phi_{k-1}^A-\phi_{k+1}^A-\big(2\phi_{k-1}^J(x)+8\phi_k^J(x)+2\phi_{k+1}^J(x)\big)\\&=\begin{cases}-2\frac{x-x_k}{h}-4\frac{x-x_{k-1}}{h}, & x\in(x_{k-1},x_k)\\-2\frac{x-x_k}{h}-4\frac{x-x_{k+1}}{h}, & x\in(x_k,x_{k+1})\\0, & \textit{otherwise.}\end{cases}\end{aligned} \quad (4.167)$$

The proof of Proposition 4.6.2 follows the one of Proposition 4.6.1, by using the definitions of the lift operators r, l, (4.161), the fact that $r([\cdot])$, $l([\cdot])\in V_h$ and the explicit form of the inverse of the matrix M_h given in Lemma 4.6.1.

We conclude that the bilinear form f_h^{β} is represented by a matrix generated by the stencil

$$r_{f_h^{\beta}}=\left(\begin{array}{cc|cc|cc}0&\frac{\beta}{h}&0&-\frac{2\beta}{h}&0&\frac{\beta}{h}\\\frac{\beta}{h}&\frac{4\beta^2-1}{2h}&-\frac{2\beta}{h}&\frac{2(4\beta^2+1)}{h}&\frac{\beta}{h}&\frac{4\beta^2-1}{2h}\end{array}\right).$$

The case $\alpha=0$ in the bilinear form $\widetilde{a}_h^{s,\alpha,\beta}$ yields the bilinear form a_h^s corresponding to the SIPG method that we already studied. Therefore, we consider directly the case $\alpha=1$, generating the LDG method. The bilinear form $\widetilde{a}_h^{s,1,\beta}$ produces a block-tri-diagonal matrix $R_h^{s,\beta}$ generated by the stencil

$$r_h^{s,\beta}=r_{b_h}+r_{c_h}+\frac{s}{h}r_{e_h}+r_{f_h^{\beta}}=\left(\begin{array}{cc|cc|cc}-\frac{1}{h}&\frac{\beta}{h}&\frac{2}{h}&-\frac{2\beta}{h}&-\frac{1}{h}&\frac{\beta}{h}\\\frac{\beta}{h}&-\frac{1}{4h}+\frac{4\beta^2-1}{2h}&-\frac{2\beta}{h}&\frac{2s-1}{2h}+\frac{2(4\beta^2+1)}{h}&\frac{\beta}{h}&-\frac{1}{4h}+\frac{4\beta^2-1}{2h}\end{array}\right).$$

We consider the following variational problem associated to the P_1-LDG semi-discretization of the wave equation:

$$\boxed{\text{Find } u_h^{s,\beta}(\cdot,t)\in V_h,\ \text{s.t. } \partial_t^2(u_h^{s,\beta}(\cdot,t),v)_{L^2(\mathbb{R})}+a_h^{s,\beta}(u_h^{s,\beta}(\cdot,t),v)=0,\ \forall v\in V_h,\ \forall t\geq 0,} \quad (4.168)$$

complemented with the initial data $u_h^{s,\beta}(\cdot,0)=u_h^0\in V_h$ and $\partial_t u_h^{s,\beta}(\cdot,0)=u_h^1\in V_h$.

By decomposing $u_h^{s,\beta}(x,t)$ in the basis $(\phi_k^A,\phi_k^J)_{k\in\mathbb{Z}}$ as $u_h^{s,\beta}(x,t)=\sum_{k\in\mathbb{Z}}\big(A_k(t)\phi_k^A(x)+J_k(t)\phi_k^J(x)\big)$, where $A_k(t)$, $J_k(t)$ denote the average and the jump of the numerical solution at the grid point

x_k and the time $t \geq 0$, we obtain that the vector $\overrightarrow{U}^h(t) = (A_k(t), J_k(t))'_{k\in\mathbb{Z}}$ is the solution of the infinite system of second-order linear ODEs:

$$M_h \partial_t^2 \overrightarrow{U}^h(t) + R_h^{s,\beta} \overrightarrow{U}^h(t) = 0, \quad \overrightarrow{U}^h(0) = \overrightarrow{U}^{h,0}, \quad \partial_t \overrightarrow{U}^h(0) = \overrightarrow{U}^{h,1}. \tag{4.169}$$

Taking SDFTs in (4.169), we obtain that the unknown $\widehat{U}^h(\xi,t) = (\widehat{A}^h(\xi,t), \widehat{J}^h(\xi,t))'$ satisfies the 2×2-system of linear ODEs depending on the parameter $\xi \in \Pi_h$:

$$M_h(\xi) \partial_t^2 \widehat{U}^h(\xi,t) + R_h^{s,\beta}(\xi) \widehat{U}^h(t) = 0, \quad \widehat{U}^h(0) = \widehat{U}^{h,0}, \quad \partial_t \widehat{U}^h(0) = \widehat{U}^{h,1}, \tag{4.170}$$

where $\widehat{M_h}(\xi)$ is the symbol of the mass matrix M_h given by (4.31) and $R_h^{s,\beta}(\xi)$ is the symbol of the stiffness matrix $R_h^{s,\beta}$, defined as

$$R_h^{s,\beta}(\xi) = \begin{pmatrix} \frac{4}{h^2}\sin^2\left(\frac{\xi h}{2}\right) & -\frac{4\beta}{h^2}\sin^2\left(\frac{\xi h}{2}\right) \\ -\frac{4\beta}{h^2}\sin^2\left(\frac{\xi h}{2}\right) & \frac{1}{h^2} r_h^{s,\beta}(\xi) \end{pmatrix},$$

with

$$r_h^{s,\beta}(\xi) = s - \cos^2\left(\frac{\xi h}{2}\right) + 2 - \cos(\xi h) + 4\beta^2(2+\cos(\xi h)).$$

The system (4.170) can be written in a simplified form as

$$\partial_t^2 \overrightarrow{U}^h(\xi,t) + S_h^{s,\beta}(\xi) \overrightarrow{U}^h(\xi,t) = 0, \tag{4.171}$$

where

$$\begin{aligned} &S_h^{s,\beta}(\xi) = M_h^{-1}(\xi) R_h^{s,\beta}(\xi) \\ &= \begin{pmatrix} \frac{4}{h^2}\sin^2\left(\frac{\xi h}{2}\right)\big(2-\cos(\xi h)+2\beta i\sin(\xi h)\big) & -\frac{4\beta}{h^2}\sin^2\left(\frac{\xi h}{2}\right)(2-\cos(\xi h)) - \frac{2i}{h^2}\sin(\xi h) r_h^{s,\beta}(\xi) \\ \frac{8}{h^2}\sin^2\left(\frac{\xi h}{2}\right)\big(i\sin(\xi h) - 2\beta(2+\cos(\xi h))\big) & -\frac{8\beta i}{h^2}\sin^2\left(\frac{\xi h}{2}\right)\sin(\xi h) + \frac{4}{h^2}(2+\cos(\xi h)) r_h^{s,\beta}(\xi) \end{pmatrix}. \end{aligned}$$

The two eigenvalues of the matrix $S_h^{s,\beta}(\xi)$ are the solutions of the quadratic equation

$$\Lambda^2 - 2\Lambda \frac{1}{h^2}\Big[2\sin^2\left(\frac{\xi h}{2}\right)(2-\cos(\xi h)) + 2(2+\cos(\xi h)) r_h^{s,\beta}(\xi)\Big] + \frac{48}{h^4}\sin^2\left(\frac{\xi h}{2}\right)\Big(r_h^{s,\beta}(\xi) - 4\beta^2\sin^2\left(\frac{\xi h}{2}\right)\Big) = 0,$$

whose solutions are a physical eigenvalue

$$\Lambda_{ph,h}^{s,\beta}(\xi) = \frac{2\sin^2\left(\frac{\xi h}{2}\right)(2-\cos(\xi h)) + 2(2+\cos(\xi h)) r_h^{s,\beta}(\xi) - \sqrt{\Delta_h^{s,\beta}(\xi)}}{h^2}$$

and a spurious one

$$\Lambda_{sp,h}^{s,\beta}(\xi) = \frac{2\sin^2\left(\frac{\xi h}{2}\right)(2-\cos(\xi h)) + 2(2+\cos(\xi h)) r_h^{s,\beta}(\xi) + \sqrt{\Delta_h^{s,\beta}(\xi)}}{h^2},$$

with $\Delta_h^{s,\beta}(\xi) = \Big[2\sin^2\left(\frac{\xi h}{2}\right)(2-\cos(\xi h)) + 2(2+\cos(\xi h)) r_h^{s,\beta}(\xi)\Big]^2 - 48\sin^2\left(\frac{\xi h}{2}\right)\Big(r_h^{s,\beta}(\xi) - 4\beta^2\sin^2\left(\frac{\xi h}{2}\right)\Big)$.

Proposition 4.6.3. *For all $(\beta \neq 0, s \geq 0)$ or $(\beta = 0, s > 0)$, the function $\Delta_h^{s,\beta}(\xi)$ is strictly positive, $\forall \xi \in \Pi_h$. For $(\beta = 0, s = 0)$, the function $\Delta_h^{s,\beta}(\xi) = 0$ iff $\xi = 0, \pm\pi/h$, otherwise it is strictly positive. Moreover, it can be written as a sum of squares of real valuated functions as follows:*

$$\Delta_h^{s,\beta}(\xi) = 4(2+\cos(\xi h))^2\Big[r_h^{s,\beta}(\xi) - \sin^2\left(\frac{\xi h}{2}\right)\frac{2+\cos^2(\xi h)}{(2+\cos(\xi h))^2}\Big]^2 + 48\sin^4\left(\frac{\xi h}{2}\right)\Big(4\beta^2 + \frac{\sin^2(\xi h)}{(2+\cos(\xi h))^2}\Big).$$

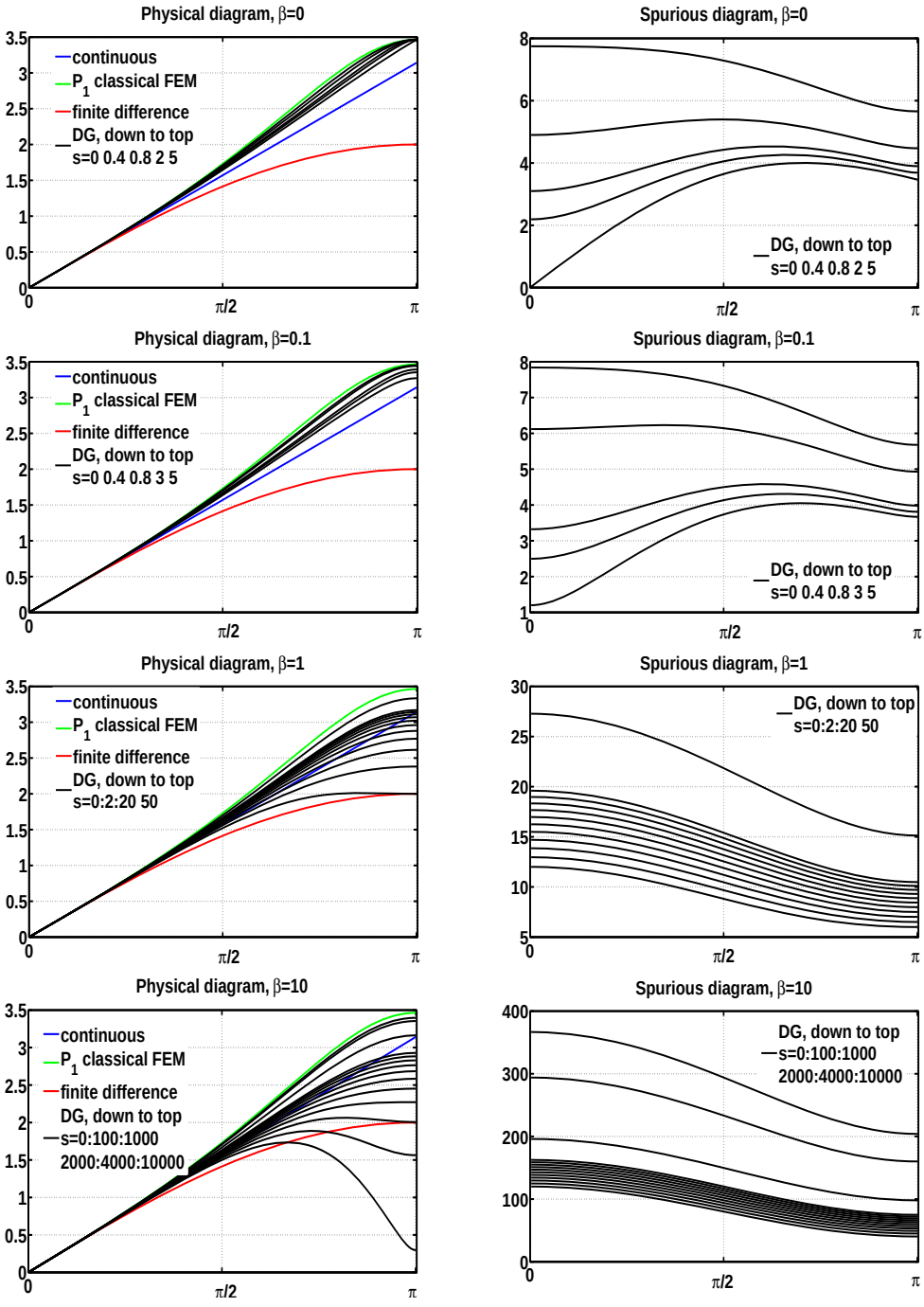

Figure 4.11: Physical and spurious dispersion relations $\lambda_{ph,1}^{s,\beta}(\xi)$ and $\lambda_{sp,1}^{s,\beta}(\xi)$, $\xi \in [0,\pi]$, for the LDG method.

In Figure 4.11, we observe that the behavior of the two dispersion relations is much more complex for the LDG methods compared to the SIPG methods, which is normal taking into account that they depend on two parameter, s and β, instead of only one as for the SIPG method. Indeed, for large values of β and small values of s, there are two critical points on the physical dispersion relation, a situation which does not appear for the SIPG method. An open problem is to find the good balance between s and β, such that the physical dispersion relation maintain only one critical point on the physical dispersion relation.

Chapter 5

Spline semi-discretizations of the $1-d$ wave equation

In this chapter, we study the behavior of the group velocity for the spline semi-discretization of the $1-d$ wave equation according to the regularity parameter $n \in \mathbb{N}$. From the analysis of some particular cases of n, we deduce that as n increases, the group velocity approximates better the continuous one. Nevertheless, as the wave number $\xi h \to \pi$, the group velocity tends to zero. We conjecture that this holds for all $n \in \mathbb{N}$. This implies that when dealing with the associated observability inequality, there are wave packets concentrated along the corresponding rays of Geometric Optics arbitrarily closed to the vertical position for which the observability constant increases at least polynomially in $1/h$ (cf. [40], [39]). For a fixed mesh size $h > 0$, we conjecture that the continuous group velocity is obtained as $n \to \infty$. However, since the support of the spline basis functions increases with n, this passing to the limit as $n \to \infty$ cannot be mimicked for the semi-discretization of the wave equation on a bounded interval.

5.1 Basic properties of B-splines.

For $n \in \mathbb{N}$, consider $P_n(a,b)$ to be the set of polynomials of order n on $(a,b) \subset \mathbb{R}$. Set $C^{-1}(\mathbb{R}) = L^\infty(\mathbb{R})$. In this chapter, we focus on the class of functions

$$\mathcal{S}_n = \{f \in C^n(\mathbb{R}) : f\chi_{(j-(n+2)/2, j-n/2)} \in P_{n+1}(j-(n+2)/2, j-n/2), \forall j \in \mathbb{Z}\}.$$

The elements of $\mathcal{S}_n$ are called cardinal spline functions of degree $n+1$ [49],[20]. The points $j-(n+2)/2$, $j \in \mathbb{Z}$, are the so-called knots of the elements of $\mathcal{S}_n$. For $n \geq -1$, let us denote by $M_n(x) \in \mathcal{S}_n$ the unique even function s.t.

$$\int_{\mathbb{R}} M_n(x)\, dx = 1.$$

Observe that $M_{-1}(x) = \chi_{[-1/2,1/2]}(x)$.

Proposition 5.1.1 (see [49], Chapters 1,2)**.** *For all $n \geq -1$, the functions $M_n(x)$ have the following properties:*

i. $supp(M_n) = [-(n+2)/2, (n+2)/2]$.

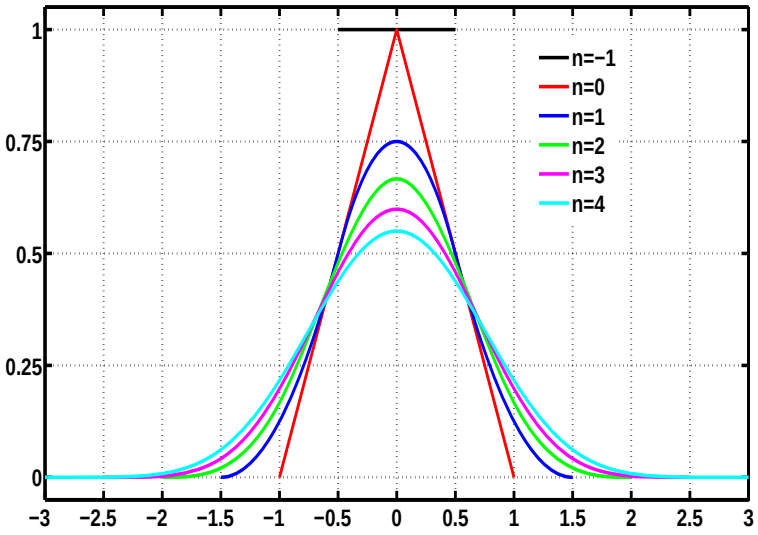

Figure 5.1: Functions $M_n(x)$, $n=-1,0,1,2,3,4$.

ii. $M_{n1+n2+2}(x)=(M_{n1}*M_{n2})(x)$, $\forall n1,n2\geq -1$. *In this way,*

$$M_{n+1}(x)=\int_{-1/2}^{1/2} M_n(x-t)\,dt.$$

iii. $\widehat{M_n}(\xi)=sinc^{n+2}(\xi/2)$, *where* $\widehat{\cdot}$ *denotes the continuous Fourier transform and* $sinc(\xi)=\sin(\xi)/\xi$ *is the Shannon function.*

iv. $(M_n(x-j))_{j\in\mathbb{Z}}$ *is an algebraic basis of* $\mathcal{S}_n$.

v. M_n *is an even function.*

vi. Under derivation, M_n *behaves as follows:*

$$M_n'(x)=M_{n-1}(x+1/2)-M_{n-1}(x-1/2),\forall n\geq 0 \tag{5.1}$$

vii. For all $n\in\mathbb{N}$, *the function* $M_n(x)$ *has the following explicit expression:*

$$M_n(x)=\frac{1}{(n+1)!}\sum_{j=0}^{n+2}(-1)^j\binom{n+2}{j}\big[\big(x+\frac{n+2}{2}-j\big)_+\big]^{n+1}, \tag{5.2}$$

where $x_+=\max(x,0)$ *is the positive part of* x.

For all $n\in\mathbb{Z}\cap[-1,\infty)$, respectively $n\in\mathbb{N}$, and for all $k\in\mathbb{Z}$, we set the following $L^2(\mathbb{R})$ and $\dot{H}^1(\mathbb{R})$-inner products:

$$m_k^n:=\int_{\mathbb{R}} M_n(x)M_n(x-k)\,dx,\quad r_k^n:=\int_{\mathbb{R}} M_n'(x)M_n'(x-k)\,dx. \tag{5.3}$$

The following result gives the explicit form of m_k^n and r_k^n for In fact, using the fact that M_n is even, we have that $m_k^n=m_{|k|}^n$ and $r_k^n=r_{|k|}^n$, so that it is sufficient to consider $k\in\mathbb{N}$.

Lemma 5.1.1. *For all $n \in \mathbb{Z} \cap [-1, \infty)$ and all $k = 0, \cdots, n+1$, the following result holds:*

$$m_k^n = M_{2n+2}(k) = \frac{1}{(2n+3)!} \sum_{j=0}^{k+n+1} (-1)^j \binom{2n+4}{j} (k+n+2-j)^{2n+3}. \tag{5.4}$$

For $k \geq n+2$,

$$m_k^n = 0. \tag{5.5}$$

For all $k \in \mathbb{N}$,

$$r_k^n = 2m_k^{n-1} - m_{k-1}^{n-1} - m_{k+1}^{n-1}. \tag{5.6}$$

Proof of Lemma 5.1.1. The identity (5.5) holds since $k \geq 0$ and for $k \geq n+2$, the intersection between the supports of $M_n(x)$ and its translation $M_n(x-k)$ has empty measure, i.e.

$$|[-(n+2)/2, (n+2)/2] \cap [k-(n+2)/2, k+(n+2)/2]| = 0.$$

The proof of (5.4) combines the properties (ii), (v) and (vii) in Proposition 5.1.1. Thus, by (ii) and (v), we obtain the first identity in (5.4). The second identity follows by (5.2) in which n is replaced by $2n+2$ and x by k. We also have to take into account the fact that $(k+n+2-j)_+ \neq 0$ for $j < k+n+2$, which concludes the proof of the second identity in (5.4).

The identity (5.6) is a consequence of the property (5.1). □

5.2 B-spline semi-discretization of the $1-d$ wave equation

Consider the Cauchy problem associated to the $1-d$ wave equation:

$$\begin{cases} \partial_t^2 u(x,t) - \partial_x^2 u(x,t) = 0, & x \in \mathbb{R}, t > 0 \\ u(x,0) = u^0(x), \partial_t u(x,0) = u^1(x), & x \in \mathbb{R}. \end{cases} \tag{5.7}$$

The variational formulation associated to this problem can be written as follows:

$$\begin{cases} u(\cdot,t) \in \dot{H}^1(\mathbb{R}) \\ \partial_t^2 \int\limits_{\mathbb{R}} u(x,t)\phi(x)\,dx + \int\limits_{\mathbb{R}} u_x(x,t)\phi_x(x,t)\,dx = 0, \forall \phi \in \dot{H}^1(\mathbb{R}). \end{cases} \tag{5.8}$$

For $h > 0$, consider an uniform grid of the real line given by the grid points $x_j = jh$, $j \in \mathbb{Z}$. In the following, for $n \in \mathbb{N}$, we will consider semi-discrete approximations of the wave equation (5.7) using spline functions of order $n+1$. For this, consider the re-scaled spline functions

$$\phi_j^n(x) = M^n\Big(\frac{x - x_j}{h}\Big). \tag{5.9}$$

According to the property (i), the functions ϕ_j^n are supported in $[x_{j-(n+2)/2}, x_{j+(n+2)/2}]$ and centered at the point x_j. Consider $V_h = \text{span}\{\phi_j^n\}_{j\in\mathbb{Z}} \subset \dot{H}^1(\mathbb{R})$.

Consider the conforming finite element method consisting in replacing in the variational formulation (5.8) the continuous solution $u(x,t)$ by a function

$$u_h^n(\cdot,t) = \sum_{j\in\mathbb{Z}} u_j^n(t)\phi_j^n(\cdot) \in V_h$$

and the test function space $\dot{H}^1(\mathbb{R})$ by V_h. Since we work on a finite-dimensional space, we can write this discrete problem in terms of the associated mass and stiffness matrices,

$$(M_h^n)_{i,j} = \int\limits_{\mathbb{R}} \phi_i^n(x)\phi_j^n(x)\,dx = hm_{|i-j|}^n, \quad (R_h^n)_{i,j} = \int\limits_{\mathbb{R}} (\phi_i^n)'(x)(\phi_j^n(x))'\,dx = r_{|i-j|}^n/h.$$

Denote by D_h the infinite Toeplitz matrix associated to the finite difference discretization of the positive definite $1-d$ Laplacian by means of the three points scheme and generated by the row

$$\left(\begin{array}{ccc} -\frac{1}{h^2} & \boxed{\frac{2}{h^2}} & -\frac{1}{h^2} \end{array} \right),$$

where the box emphasizes the diagonal element. Set M^n to be the Toeplitz matrix generated by the row

$$\left(\begin{array}{ccccccccccc} m_{n+1}^n & m_n^n & \cdots & m_2^n & m_1^n & \boxed{m_0^n} & m_1^n & m_2^n & \cdots & m_n^n & m_{n+1}^n \end{array} \right).$$

In Table 5.2, we exemplify the row generating the matrix M^n, for $n = -1, 0, 1, 2, 3, 4$. Let us

n	m_0^n	m_1^n	m_2^n	m_3^n	m_4^n	m_5^n
-1	1	0	0	0	0	0
0	$\frac{4}{6}$	$\frac{1}{6}$	0	0	0	0
1	$\frac{66}{120}$	$\frac{26}{120}$	$\frac{1}{120}$	0	0	0
2	$\frac{2416}{5040}$	$\frac{1191}{5040}$	$\frac{120}{5040}$	$\frac{1}{5040}$	0	0
3	$\frac{156190}{362880}$	$\frac{88234}{362880}$	$\frac{14608}{362880}$	$\frac{502}{362880}$	$\frac{1}{362880}$	0
4	$\frac{15724248}{39916800}$	$\frac{9738114}{39916800}$	$\frac{2203488}{39916800}$	$\frac{152637}{39916800}$	$\frac{2036}{39916800}$	$\frac{1}{39916800}$

Table 5.1: The elements of the row generating the matrix M^n for $n = -1, 0, 1, 2, 3, 4$.

observe that for any $n \in \mathbb{Z} \cap [-1, \infty)$, the following property holds:

$$\sum_{k=-(n+1)}^{n+1} m_k^n = 1. \tag{5.10}$$

This can be obtained by an inductive argument, using the property (ii) in Proposition 5.1.1.

Denoting by $\overrightarrow{U}^n(t) = (u_j^n(t))_{j\in\mathbb{Z}}$ the unknown vector and taking into account Lemma 5.1.1, it is easy to see that $\overrightarrow{U}^n(t)$ verifies the following Cauchy problem associated to the linear system of second order ODEs:

$$\left\{ \begin{array}{l} M^n \partial_t^2 \overrightarrow{U}^n(t) + D_h M^{n-1} \overrightarrow{U}^n(t) = 0, \quad t > 0 \\ \overrightarrow{U}^n(0) = \overrightarrow{U}^{n,0},\ \partial_t \overrightarrow{U}^n(0) = \overrightarrow{U}^{n,1}, \end{array} \right. \tag{5.11}$$

where the vectors $\overrightarrow{U}^{n,i}$, $i = 0, 1$, are obtained, for example, from the initial data (u^0, u^1) in (5.7) by solving the equations $M_h^n \overrightarrow{U}^{n,i} = \overrightarrow{F}^i$, with

$$F_j^i = \int_{\mathbb{R}} u^i(x) \phi_j^n(x)\, dx.$$

Fourier analysis of the spline semi-discretization of the $1-d$ wave equation. Set $\Pi_h = [-\pi/h, \pi/h]$. Denote by $\widehat{u}^{h,n}(\xi, t)$ the SDFT at the scale h of the sequence of unknowns $\overrightarrow{U}^n(t)$ in (5.11). By applying the SDFT to (5.11), we obtain

$$\left\{ \begin{array}{ll} m^{h,n}(\xi) \partial_t^2 \widehat{u}^{h,n}(\xi, t) + \Omega^h(\xi) m^{h,n-1}(\xi) \widehat{u}^{h,n}(\xi) = 0, & \xi \in \Pi_h, t > 0 \\ \widehat{u}^{h,n}(\xi, 0) = \widehat{u}_0^h(\xi),\ \partial_t \widehat{u}^{h,n}(\xi, 0) = \widehat{u}_1^h(\xi), & \xi \in \Pi_h \end{array} \right. \tag{5.12}$$

where $m^{h,n}(\xi)$ and $\Omega^h(\xi)$ are the Fourier symbols of the matrices M^n and D_h.

Lemma 5.2.1. *The Fourier symbols $\widehat{m}^{h,n}(\xi)$ and $\widehat{\Omega}^h(\xi)$ of the matrices M^n and D_h are explicitly given as follows:*

$$m^{h,n}(\xi) = m_0^n + 2\sum_{k=1}^{n+1} m_k^n \cos(k\xi h), \quad \Omega^h(\xi) = \frac{4}{h^2}\sin^2\Big(\frac{\xi h}{2}\Big). \tag{5.13}$$

Proof of Lemma 5.2.1. We do not insist on the explicit form of $\Omega^h(\xi)$, since D_h is the most classical matrix arising in the approximation of the $1-d$ positive Laplacian by finite differences. The symbol $m^{h,n}(\xi)$ is by definition a function such that for any sequence $\overrightarrow{f}$, $\widehat{M^n f}(\xi) = m^{h,n}(\xi)\widehat{f}^h(\xi)$, with $\widehat{f}^h(\xi)$ being the SDFT of the sequence $\overrightarrow{f}$. Thus, we choose a sequence $\overrightarrow{f}$ and we compute the SDFT of $M^n\overrightarrow{f}$, whose terms are explicitly given by

$$(M^n\overrightarrow{f})_j = \sum_{k=-(n+1)}^{n+1} m_{|k|}^n f_{j+k}.$$

Therefore,

$$\begin{aligned}\widehat{M^n f}^h(\xi) &= h\sum_{j\in\mathbb{Z}}(M^n\overrightarrow{f})_j \exp(-i\xi x_j) = \sum_{k=-(n+1)}^{n+1} m_{|k|}^n\Big(h\sum_{j\in\mathbb{Z}} f_{j+k}\exp(-i\xi x_{j+k})\Big)\exp(i\xi x_k)\\ &= \widehat{f}^h(\xi)\sum_{k=-(n+1)}^{n+1} m_{|k|}^n \exp(i\xi x_k),\end{aligned}$$

which concludes the explicit form of $m^{h,n}(\xi)$ given by (5.13). □

n	$m^{h,n}(\xi)$
-1	1
0	$\frac{1}{3}\big(1+2\theta\big)$
1	$\frac{1}{15}\big(2+11\theta+2\theta^2\big)$
2	$\frac{1}{315}\big(17+180\theta+114\theta^2+4\theta^3\big)$
3	$\frac{1}{2835}\big(62+1072\theta+1452\theta^2+247\theta^3+2\theta^4\big)$
4	$\frac{1}{155925}\big(1382+35396\theta+8302\theta^2+34096\theta^3+2026\theta^4+4\theta^5\big)$

Table 5.2: The Fourier symbols $m^{h,n}(\xi)$ for $n=-1,0,1,2,3,4$, with $\theta=\cos^2\big(\frac{\xi h}{2}\big)$.

In Table 5.2, we observe that for the particular cases $-1\le n\le 4$, $m^{h,n}(\xi)>0$, $\forall\xi\in\Pi_h$.

Conjecture 5.2.1. *For all $n\in\mathbb{Z}\cap[-1,\infty)$, there exists a vector depending on n, $(a_j^n)_{j=0,\cdots,n+1}$, such that $a_j^n>0$ and*

$$m^{h,n}(\xi) = \sum_{j=0}^{n+1} a_j^n\theta^j,\ \theta=\cos^2\Big(\frac{\xi h}{2}\Big). \tag{5.14}$$

For $\xi\in\Pi_1$, set $m^n(\xi) := m^{1,n}(\xi)$. Then $m^{h,n}(\xi) = m^n(\xi h)$, for all $\xi\in\Pi_h$. If Conjecture 5.2.1 would be true, then the dispersion relation associated to the spline semi-discretization (5.11), given by

$$\lambda^{h,n}(\xi) = \frac{2}{h}\sin\Big(\frac{\xi h}{2}\Big)\sqrt{\frac{m^{n-1}(\xi h)}{m^n(\xi h)}}, \tag{5.15}$$

is purely real. However, observe that, due to the property (5.10), for any $\xi \in \Pi_h$, $\xi = O(1)$ as $h \to 0$, we get $\lim_{h\to 0} \lambda^{h,n}(\xi) = \xi$, meaning that the spline approximation is spectrally correct, in the sense that the low frequencies converge to the continuous ones as $h \to 0$.

Let us introduce the following two functions:

$$\partial_\xi m^n(\xi h) = -h\cos\Big(\frac{\xi h}{2}\Big)\sin\Big(\frac{\xi h}{2}\Big)d^n(\xi h),\ d^n(\xi h) := \sum_{j=1}^{n+1} a_j^n j\theta^{j-1},\ \theta = \cos^2\Big(\frac{\xi h}{2}\Big) \tag{5.16}$$

and for $n \in \mathbb{N}$

$$e^n(\xi h) = m^n(\xi h)m^{n-1}(\xi h) + \sin^2\Big(\frac{\xi h}{2}\Big)\big(-d^{n-1}(\xi h)m^n(\xi h) + d^n(\xi h)m^{n-1}(\xi h)\big). \tag{5.17}$$

Conjecture 5.2.2. *For all $\xi \in \Pi_h$ and $n \in \mathbb{N}$, $\lim_{\xi h\to 0} e^n(\xi h) = 1$ and $e^n(\xi h) > 0$.*

n	$e^n(\xi h)$
0	1
1	$\frac{1}{15}\big(3+4\theta+8\theta^2\big)$
2	$\frac{1}{1575}\big(69+254\theta+924\theta^2+264\theta^3+64\theta^4\big)$
3	$\frac{1}{945^2}\big(8118+57552\theta+340485\theta^2+317490\theta^3+157260\theta^4+11352\theta^5+768\theta^6\big)$
4	$\frac{1}{55\cdot 2835^2}\big(798732+9244284\theta+81738516\theta^2+1560350880\theta^3+$ $+1451666850\theta^4+425193780\theta^5+64148760\theta^6+1267440\theta^7+30720\theta^8\big)$

Table 5.3: The Fourier symbols $e^n(\xi h)$ for $n = 0, 1, 2, 3, 4$, with $\theta = \cos^2\big(\frac{\xi h}{2}\big)$.

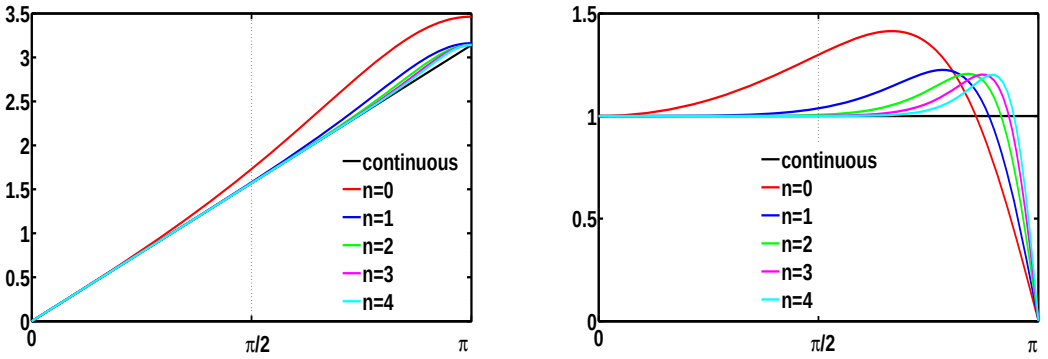

Figure 5.2: Dispersion relation $\lambda^{1,n}(\xi)$ versus group velocity $\partial_\xi\lambda^{1,n}(\xi)$ for the spline semi-discretization of the wave equation for $n = 0, 1, 2, 3, 4$.

Lemma 5.2.2. *For all $n \in \mathbb{N}$, the group velocity $\partial_\xi\lambda^{h,n}(\xi)$ has the following explicit form:*

$$\partial_\xi\lambda^{h,n}(\xi) = \frac{\cos\Big(\frac{\xi h}{2}\Big)e^n(\xi h)}{(m^n(\xi h))^{3/2}(m^{n-1}(\xi h))^{1/2}}. \tag{5.18}$$

Moreover, $\partial_\xi\lambda^{h,n}(\xi) = 0$ if and only if $\xi h = \pm\pi$ and $\partial_\xi\lambda^{h,n}(\xi) > 0$ for all $\xi \in \Pi_h \setminus \{\pm\pi/h\}$.

Conjecture 5.2.3 (The behavior of the group velocity $\partial_\xi \lambda^{h,n}(\xi)$ as $n \to \infty$)**.** *For a fixed $h > 0$ and for all $\xi \in \Pi_h$, the following two properties hold*

i. There exists $e(\xi h) = \lim_{n\to\infty} e^n(\xi h) > 0$ $(= 1)$, for all $\xi \in \Pi_h$.

ii. Consider the vector $(b_j^n)_{j=0,4n+3}$ such that $(m^n(\xi h))^3 m^{n-1}(\xi h) = \sum_{j=0}^{4n+3} b_j^n \theta^j$, with $\theta = \cos^2(\xi h/2)$. Then $b_0^n \to 0$ as $n \to \infty$ and there exists $m(\xi h) = \lim_{n\to\infty} b_j^n \theta^{j-1} (= 1)$, for all $\xi \in \Pi_h$.

In this way, for all $\xi \in \Pi_h$, $\lim_{n\to\infty} \partial_\xi \lambda^{h,n}(\xi) = 1$.

This means that, in the limit as $n \to \infty$, the continuous group velocity is recovered and once with that the uniform observability property.

Chapter 6

Higher-order classical finite element semi-discretizations of the $1-d$ wave equation

This chapter deals with the quadratic classical finite element semi-discretization of the wave equation. A similar analysis was done for some classes of DG methods in [41] and in the second chapter. Firstly, a carefully Fourier analysis of this method is carried on. Two dispersion branches appear: one of physical nature, the so-called *acoustical dispersion*, and one of spurious type, the so-called *optical dispersion relation* (for more details, concerning this terminology, see [29], [15]). In fact, in [29], the explicit form of the two dispersion curves was indicated. Moreover, we study the two group velocities associated to this approximation and we see that some pathological behavior has to be expected concerning the propagation properties. Thus, there are three singular points located on the two dispersion diagrams: one on the acoustical dispersion at the wave number $\xi = \pi/h$ and two on the optical dispersion located at the wave numbers $\xi = 0$ and $\xi = \pi/h$. Consequently, there are wave packets propagating with an arbitrarily small speed located on each one of the two diagrams, neglecting in this way the uniform observability property (cf. [41]).

In the second part of the chapter, we deal with some filtering mechanisms, i.e. we construct classes of initial data within which the propagation properties of the continuous model are recovered uniformly in the mesh size parameter h. Four filtering mechanisms are analyzed:

A. Concentration on the acoustical mode + Fourier filtering. Both initial position and velocity are chosen in the Fourier space as the eigenvector corresponding to the acoustical branch times a certain profile whose support is included in $\Pi_h^\delta = [-\pi\delta/h, \pi\delta/h]$, for some $\delta \in (0,1)$.

B. Concentration on the acoustical mode + bi-grid algorithm. Both initial position and velocity are chosen in the Fourier space as the eigenvector corresponding to the acoustical branch times a certain profile which is given by a bi-grid algorithm, i.e. by linear interpolation from a coarse uniform grid of size $2h$ to a finer uniform grid of size h. A dyadic decomposition argument like in [32] or [41] is used. In order to do this, the total energy of initial data in this class is proved to be bounded by the energy of their projection on the first half of the spectrum.

C. Linear initial data + Fourier filtering on the nodal components. The value corresponding to the midpoints in the initial data are given by linear interpolation between the values corresponding to the two neighboring nodal points. Moreover, the Fourier transforms of the sequences of nodal values are supported in Π_h^δ.

D. Linear initial data + bi-grid algorithm on the nodal components. With respect to the previous filtering mechanism, the values corresponding to the nodal points in the initial data are given by a bi-grid algorithm. This is the most important filtering mechanism from an

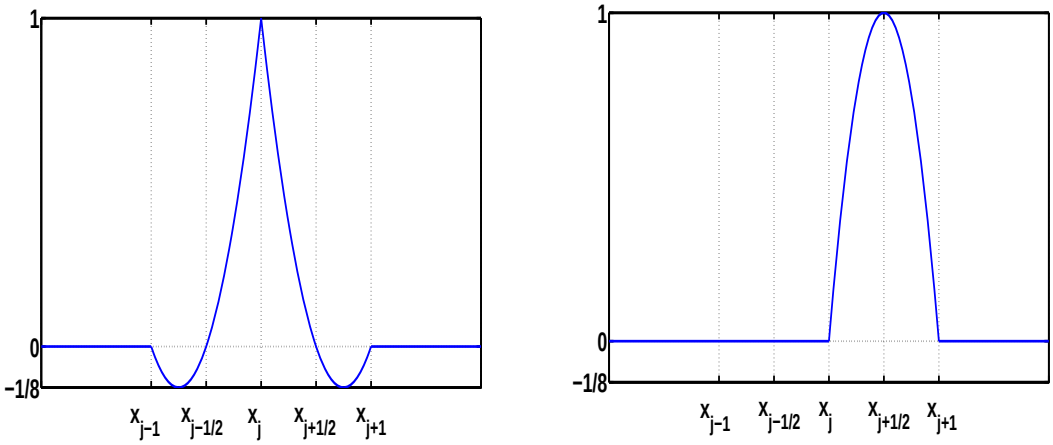

Figure 6.1: The two basis functions: ϕ_j^N (left) and $\phi_{j+1/2}^H$ (right).

applicability point of view, since it can be implemented without using the SDFT.

6.1 Fourier analysis of the quadratic finite element method

Consider an uniform grid of mesh size $h > 0$ of the whole real line constituted by the grid points $x_j = jh, j \in \mathbb{Z}$. In this section, we deal with a conforming approximation of the wave equation (5.7), but the space V_h is generated by basis functions of two types, instead of being generated by functions of only one type as in the third chapter. This time,

$$V_h = \{\phi \in \dot{H}^1(\mathbb{R}) : \phi|_{x_j}^{x_{j+1}} \in P_2\}.$$

Observe that, since $V_h \subset \dot{H}^1(\mathbb{R})$, its element are continuous functions at the grid points. A polynomial of order two in each interval $[x_j, x_{j+1}]$ is completely determined given its value at three points. Two of them are the two endpoints. We choose the third one to be the midpoint $x_{j+1/2}$. Then, the space V_h is generated by two kinds of functions: one designed to represent the nodal values, called ϕ_j^N, and the other one dealing with the midpoint values, called $\phi_{j+1/2}^H$.

The solution of the associated semi-discrete $1-d$ wave equation can be then represented according to the basis functions as follows:

$$u_h(x,t) = \sum_{j\in\mathbb{Z}} N_j(t)\phi_j^N(x) + \sum_{j\in\mathbb{Z}} H_{j+1/2}(t)\phi_{j+1/2}^H(x). \tag{6.1}$$

Set $\overrightarrow{U}^h(t) := (N_j(t), H_{j+1/2}(t))_{j\in\mathbb{Z}}$. This vector is the unknown in the following homogeneous system of linear second-order differential equations:

$$\begin{cases} M_h\partial_t^2\overrightarrow{U}^h(t) + R_h\overrightarrow{U}^h(t) = 0, & t > 0 \\ \overrightarrow{U}^h(0) = \overrightarrow{U}^{h,0}, \partial_t\overrightarrow{U}^h(0) = \overrightarrow{U}^{h,1}, \end{cases} \tag{6.2}$$

with M_h and R_h being infinite mass and stiffness pentha-diagonal matrices generated by the stencils:

$$m_h = \begin{pmatrix} -\frac{h}{30} & \frac{h}{15} & \frac{4h}{15} & \frac{h}{15} & -\frac{h}{30} \\ 0 & 0 & \frac{h}{15} & \frac{8h}{15} & \frac{h}{15} \end{pmatrix}, \quad r_h = \begin{pmatrix} \frac{1}{3h} & -\frac{8}{3h} & \frac{14}{3h} & -\frac{8}{3h} & \frac{1}{3h} \\ 0 & 0 & -\frac{8}{3h} & \frac{16}{3h} & -\frac{8}{3h} \end{pmatrix}. \tag{6.3}$$

System (6.2) is in fact an infinite system of two kinds of difference equations. We can write it in the following more explicit way

$$\begin{cases} m_h\big(\partial_t^2 N_{j-1}(t), \partial_t^2 H_{j-1/2}(t), \partial_t^2 N_j(t), \partial_t^2 H_{j+12/2}(t), \partial_t^2 N_{j+1}(t)\big)' \\ \qquad +r_h\big(N_{j-1}(t), H_{j-1/2}(t), N_j(t), H_{j+12/2}(t), N_{j+1}(t)\big)' = (0,0)', \quad t>0, j\in\mathbb{Z}, \\ \overrightarrow{U}^h(0) = \overrightarrow{U}^{h,0}, \partial_t \overrightarrow{U}^h(0) = \overrightarrow{U}^{h,1}. \end{cases}$$

Set $\Pi_h := [-\pi/h, \pi/h]$ and $\widehat{N}^h(\xi,t)$, $\widehat{H}^h(\xi,t)$ to be the SDFTs (for the precise definition, see (3.13) in Chapter 1) of the sequences of unknowns $\overrightarrow{N}(t) = (N_j(t))_{j\in\mathbb{Z}}$ and $\overrightarrow{H}(t) = (H_{j+1/2}(t))_{j\in\mathbb{Z}}$ and $\widehat{U}^h(\xi,t) = (\widehat{N}^h(\xi,t), \widehat{H}^h(\xi,t))'$. By taking the SDFT in (6.2), we obtain

$$\begin{cases} \partial_t^2 \widehat{U}^h(\xi,t) + A_h(\xi)\widehat{U}^h(\xi,t) = 0, & \xi\in\Pi_h, t>0 \\ \widehat{U}^h(\xi,0) = \widehat{U}^{h,0}(\xi), \quad \partial_t \widehat{U}^h(\xi,0) = \widehat{U}^{h,1}(\xi), & \xi\in\Pi_h, \end{cases} \tag{6.4}$$

where

$$A_h(\xi) := (M_h(\xi))^{-1} R_h(\xi) = \begin{pmatrix} \frac{8\left(3+2\cos^2\left(\frac{\xi h}{2}\right)\right)}{h^2\left(1+\sin^2\left(\frac{\xi h}{2}\right)\right)} & -\frac{40\cos\left(\frac{\xi h}{2}\right)}{h^2\left(1+\sin^2\left(\frac{\xi h}{2}\right)\right)} \\ -\frac{2\cos\left(\frac{\xi h}{2}\right)\left(10+3\sin^2\left(\frac{\xi h}{2}\right)\right)}{h^2\left(1+\sin^2\left(\frac{\xi h}{2}\right)\right)} & \frac{20}{h^2\left(1+\sin^2\left(\frac{\xi h}{2}\right)\right)} \end{pmatrix} \tag{6.5}$$

and $M_h(\xi)$ and $R_h(\xi)$ are the 2×2 symbol matrices corresponding to the mass and stiffness matrices, M_h and R_h, given by

$$M_h(\xi) = \begin{pmatrix} \frac{4-\cos(\xi h)}{15} & \frac{2}{15}\cos\left(\frac{\xi h}{2}\right) \\ \frac{2}{15}\cos\left(\frac{\xi h}{2}\right) & \frac{8}{15} \end{pmatrix}, \; R_h(\xi) = \begin{pmatrix} \frac{14+2\cos(\xi h)}{3h^2} & -\frac{16}{3h^2}\cos\left(\frac{\xi h}{2}\right) \\ -\frac{16}{3h^2}\cos\left(\frac{\xi h}{2}\right) & \frac{16}{3h^2} \end{pmatrix}.$$

Properties of the eigenvalues and eigenvectors of $A_h(\xi)$. Set

$$\Delta_h(\xi) := 1 + 268\cos^2\big(\frac{\xi h}{2}\big) - 44\cos^4\big(\frac{\xi h}{2}\big). \tag{6.6}$$

The eigenvalues of the matrix $A_h(\xi)$, solutions of the second-order characteristic equation

$$\Big(\Lambda - \frac{8\big(3+2\cos^2\left(\frac{\xi h}{2}\right)\big)}{h^2\big(1+\sin^2\left(\frac{\xi h}{2}\right)\big)}\Big)\Big(\Lambda - \frac{20}{h^2\big(1+\sin^2\left(\frac{\xi h}{2}\right)\big)}\Big) - \frac{80\cos^2\left(\frac{\xi h}{2}\right)\big(10+3\sin^2\left(\frac{\xi h}{2}\right)\big)}{h^4\big(1+\sin^2\left(\frac{\xi h}{2}\right)\big)^2} = 0, \tag{6.7}$$

are given explicitly by

$$\Lambda_h^a(\xi) = \frac{120\sin^2\left(\frac{\xi h}{2}\right)}{h^2\big(11+4\cos^2\left(\frac{\xi h}{2}\right)+\sqrt{\Delta_h(\xi)}\big)}, \quad \Lambda_h^o(\xi) = \frac{22+8\cos^2\left(\frac{\xi h}{2}\right)+2\sqrt{\Delta_h(\xi)}}{h^2\big(1+\sin^2\left(\frac{\xi h}{2}\right)\big)}, \tag{6.8}$$

the superscripts "a" and "o" standing for "acoustical" and "optical". Denote by $\lambda_h^a(\xi)$ and $\lambda_h^o(\xi)$ the square root of $\Lambda_h^a(\xi)$ and, respectively, $\Lambda_h^o(\xi)$ for $\xi\in[0,\pi/h]$ and their odd extension for $\xi\in[-\pi/h,0]$.

The two eigenvectors corresponding to the acoustical and optical branches $\Lambda_h^a(\xi)$ and $\Lambda_h^o(\xi)$ are

$$P_h^a(\xi) = \begin{pmatrix} p_{h,1}^a(\xi) \\ p_{h,2}^a(\xi) \end{pmatrix} = \begin{pmatrix} \frac{20\cos\left(\frac{\xi h}{2}\right)}{\sqrt{400\cos^2\left(\frac{\xi h}{2}\right)+\left(1+4\cos^2\left(\frac{\xi h}{2}\right)+\sqrt{\Delta_h(\xi)}\right)^2}} \\ \frac{1+4\cos^2\left(\frac{\xi h}{2}\right)+\sqrt{\Delta_h(\xi)}}{\sqrt{400\cos^2\left(\frac{\xi h}{2}\right)+\left(1+4\cos^2\left(\frac{\xi h}{2}\right)+\sqrt{\Delta_h(\xi)}\right)^2}} \end{pmatrix} \tag{6.9}$$

and, respectively,

$$P_h^o(\xi) = \begin{pmatrix} p_{h,1}^o(\xi) \\ p_{h,2}^o(\xi) \end{pmatrix} = \begin{pmatrix} \frac{1+4\cos^2\left(\frac{\xi h}{2}\right)+\sqrt{\Delta_h(\xi)}}{\sqrt{\cos^2\left(\frac{\xi h}{2}\right)\left(10+3\sin^2\left(\frac{\xi h}{2}\right)\right)^2+\left(1+4\cos^2\left(\frac{\xi h}{2}\right)+\sqrt{\Delta_h(\xi)}\right)^2}} \\ -\frac{\cos\left(\frac{\xi h}{2}\right)\left(10+3\sin^2\left(\frac{\xi h}{2}\right)\right)}{\sqrt{\cos^2\left(\frac{\xi h}{2}\right)\left(10+3\sin^2\left(\frac{\xi h}{2}\right)\right)^2+\left(1+4\cos^2\left(\frac{\xi h}{2}\right)+\sqrt{\Delta_h(\xi)}\right)^2}} \end{pmatrix}. \tag{6.10}$$

Proposition 6.1.1 (Properties of the acoustical branch)**.** *The acoustical eigenvalue $\Lambda_h^a(\xi)$ and the corresponding eigenvector $P_h^a(\xi)$ verify the following properties:*

A1. $\lim_{\xi\to 0}\lambda_h^a(\xi) = 0$, $\lim_{\xi\to\pi/h}\lambda_h^a(\xi) = \frac{\sqrt{10}}{h}$.

A2. $\lim_{\xi\to 0} P_h^a(\xi) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})'$, $\lim_{\xi\to\pi/h} P_h^a(\xi) = (0,1)'$.

A3. $\lim_{\xi\to 0}\partial_\xi\lambda_h^a(\xi) = 1$, $\lim_{\xi\to\pi/h}\partial_\xi\lambda_h^a(\xi) = 0$, $\partial_\xi\lambda_h^a(\xi) > 0$, $\forall\xi\in(0,\pi/h)$.

Proof. The first two properties (A1) and (A2) are obvious. In order to prove (A3), we firstly compute the group velocity corresponding to the acoustical dispersion relation

$$\partial_\xi\lambda_h^a(\xi) = \frac{15\sqrt{30}\cos\left(\frac{\xi h}{2}\right)}{\left(11+4\cos^2\left(\frac{\xi h}{2}\right)+\sqrt{\Delta_h(\xi)}\right)^{3/2}}\frac{1}{\sqrt{\Delta_h(\xi)}}\Big(9+6\cos^2\left(\frac{\xi h}{2}\right)+\sqrt{\Delta_h(\xi)}\Big). \tag{6.11}$$

From this explicit expression, we see that the group velocity is zero iff $\cos(\xi h/2) = 0$, i.e. $\xi = \pi/h$. For the other cases, $\xi\in(0,\pi/h)$, $\partial_\xi\lambda_h^a(\xi) > 0$, i.e. $\lambda_h^a(\xi)$ is strictly increasing with respect to ξ. □

Proposition 6.1.2 (Properties of the optical branch)**.** *The acoustical eigenvalue $\Lambda_h^o(\xi)$ and the corresponding eigenvector $P_h^o(\xi)$ verify the following properties:*

O1. $\lim_{\xi\to 0}\lambda_h^o(\xi) = \sqrt{60}/h$, $\lim_{\xi\to\pi/h}\lambda_h^o(\xi) = \frac{\sqrt{12}}{h}$.

O2. $\lim_{\xi\to 0} P_h^o(\xi) = (\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}})'$, $\lim_{\xi\to\pi/h} P_h^o(\xi) = (1,0)'$.

O3. $\lim_{\xi\to 0}\partial_\xi\lambda_h^o(\xi) = \lim_{\xi\to\pi/h}\partial_\xi\lambda_h^o(\xi) = 0$, $\partial_\xi\lambda_h^o(\xi) < 0$, $\forall\xi\in(0,\pi/h)$.

Proof. The first two properties (O1) and (O2) are obvious. The third one, (O3), is obtained from the explicit expression of the group velocity corresponding to the optical dispersion relation,

$$\partial_\xi\lambda_h^o(\xi) = -\frac{1}{\sqrt{\Delta_h(\xi)}}\frac{\sin\left(\frac{\xi h}{2}\right)\cos\left(\frac{\xi h}{2}\right)\Big(269+46\cos^2\left(\frac{\xi h}{2}\right)+19\sqrt{\Delta_h(\xi)}\Big)}{\left(1+\sin^2\left(\frac{\xi h}{2}\right)\right)^{3/2}\Big(22+8\cos^2\left(\frac{\xi h}{2}\right)+2\sqrt{\Delta_h(\xi)}\Big)^{1/2}}. \tag{6.12}$$

□

From this explicit expression, we observe that $\partial_\xi\lambda_h^o(\xi) = 0$ iff $\sin(\xi h/2)\cos(\xi h/2) = 0$, i.e. $\xi = 0$ or $\xi = \pi/h$. For the other cases of $\xi\in(0,\pi/h)$, we have $\partial_\xi\lambda_h^o(\xi) < 0$, i.e. $\lambda_h^o(\xi)$ is strictly decreasing in $\xi\in(0,\pi/h)$.

The matrix $A_h(\xi)$ can be decomposed as $A_h(\xi) = P_h(\xi)\Lambda_h(\xi)(P_h(\xi))^{-1}$, where

$$\Lambda_h(\xi) = \begin{pmatrix} \Lambda_h^a(\xi) & 0 \\ 0 & \Lambda_h^o(\xi) \end{pmatrix}, \quad P_h(\xi) = \begin{pmatrix} p_{h,1}^a(\xi) & p_{h,1}^o(\xi) \\ p_{h,2}^a(\xi) & p_{h,2}^o(\xi) \end{pmatrix} \tag{6.13}$$

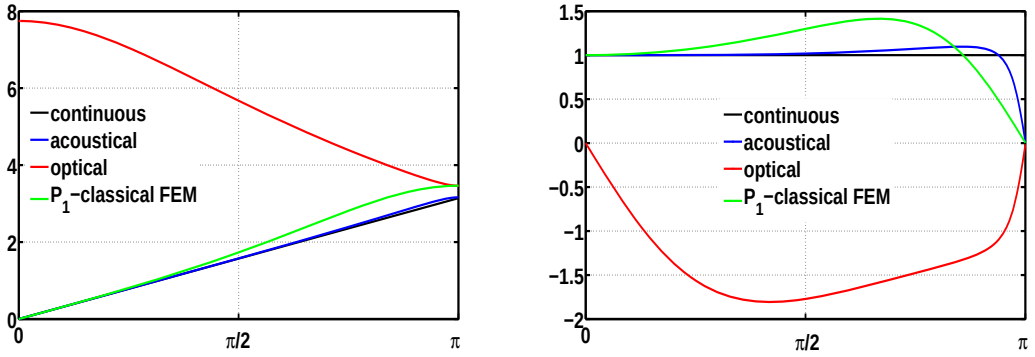

Figure 6.2: Dispersion relations versus group velocities for the P_2-classical finite element semi-discretization of the wave equation

and $(P_h(\xi))^{-1}$ is the inverse of the matrix $P_h(\xi)$.

The solution of the equation (6.4) is given by

$$\widehat{U}^h(\xi,t) = \sum_{\pm}\frac{1}{2}\Big[P_h(\xi)\begin{pmatrix}\exp(\pm it\lambda_h^a(\xi)) & 0\\ 0 & \exp(\pm it\lambda_h^o(\xi))\end{pmatrix}(P_h(\xi))^{-1}\widehat{U}^{h,0}(\xi) \tag{6.14}$$
$$+ P_h(\xi)\begin{pmatrix}\pm\frac{\exp(\pm it\lambda_h^a(\xi))}{i\lambda_h^a(\xi)} & 0\\ 0 & \pm\frac{\exp(\pm it\lambda_h^o(\xi))}{i\lambda_h^o(\xi)}\end{pmatrix}(P_h(\xi))^{-1}\widehat{U}^{h,1}(\xi)\Big].$$

Consider $O\subset\mathbb{R}$. As for the DG semi-disretzation, we divide the total energy of the solution $\overrightarrow{U}^h(t)=(N_j(t),H_{j+1/2}(t))'$ of (6.2) concentrated in the set O at time t as follows:

$$E_{O,h}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\sum_{k=1}^{8}E_{O,h}^k(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t),$$

where

$$E_{O,h}^1(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{h}{60}\sum_{x_j\in O}\big[|\partial_t N_j(t)+\partial_t H_{j+1/2}(t)|^2+|\partial_t N_{j-1}(t)+\partial_t H_{j-1/2}(t)|^2\big],$$

$$E_{O,h}^2(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{h}{60}\sum_{x_j\in O}\big[|\partial_t H_{j+1/2}(t)+\partial_t N_{j+1}(t)|^2+|\partial_t H_{j-1/2}(t)+\partial_t N_j(t)|^2\big],$$

$$E_{O,h}^3(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{h}{120}\sum_{x_j\in O}\big[|\partial_t N_{j+1}(t)-\partial_t N_j(t)|^2+|\partial_t N_j(t)-\partial_t N_{j-1}(t)|^2\big],$$

$$E_{O,h}^4(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{h}{30}\sum_{x_j\in O}|\partial_t N_j(t)|^2,$$

$$E_{O,h}^5(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{h}{10}\sum_{x_j\in O}\big[|\partial_t H_{j+1/2}(t)|^2+|\partial_t H_{j-1/2}(t)|^2\big],$$

$$E_{O,h}^6(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{h}{24}\sum_{x_j\in O}\Big[\Big|\frac{3N_j(t)-4H_{j+1/2}(t)+N_{j+1}(t)}{h}\Big|^2+\Big|\frac{3N_{j-1}(t)-4H_{j-1/2}(t)+N_j(t)}{h}\Big|^2\Big],$$

$$E_{O,h}^7(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{h}{24}\sum_{x_j\in O}\Big[\Big|\frac{N_j(t)-4H_{j+1/2}(t)+3N_{j+1}(t)}{h}\Big|^2+\Big|\frac{N_{j-1}(t)-4H_{j-1/2}(t)+3N_j(t)}{h}\Big|^2\Big]$$

and

$$E^8_{O,h}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)=\frac{h}{6}\sum_{x_j\in O}\Big[\Big|\frac{N_{j+1}(t)-N_j(t)}{h}\Big|^2+\Big|\frac{N_j(t)-N_{j-1}(t)}{h}\Big|^2\Big].$$

Define the total energy of the system to be

$$E_h(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1}):=E_{\mathbb{R},h}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)$$

and remark that the total energy is conserved in time.

The Fourier representation of the total energy is

$$\begin{aligned}E_h(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1})&=\frac{1}{4\pi}\int\limits_{\Pi_h}\big(\langle M_h(\xi)\overrightarrow{U}^{h,1}(\xi),\overrightarrow{U}^{h,1}(\xi)\rangle+\langle R_h(\xi)\overrightarrow{U}^{h,0}(\xi),\overrightarrow{U}^{h,0}(\xi)\rangle\big)\,d\xi \qquad (6.15)\\
&=\frac{1}{30\pi}\int\limits_{\Pi_h}|\widehat{N}^{h,1}(\xi)+\cos(\xi h/2)\widehat{H}^{h,1}(\xi)|^2\,d\xi+\frac{1}{60\pi}\int\limits_{\Pi_h}\big(1+2\sin^2(\xi h/2)\big)|\widehat{N}^{h,1}(\xi)|^2\,d\xi\\
&+\frac{1}{30\pi}\int\limits_{\Pi_h}\big(3+\sin^2(\xi h/2)\big)|\widehat{H}^{h,1}(\xi)|^2\,d\xi+\frac{4}{3\pi h^2}\int\limits_{\Pi_h}\big|\cos(\xi h/2)\widehat{N}^{h,0}(\xi)-\widehat{H}^{h,0}(\xi)\big|^2\,d\xi\\
&+\frac{1}{4\pi}\int\limits_{\Pi_h}\frac{4}{h^2}\sin^2(\xi h/2)|\widehat{N}^{h,0}(\xi)|^2\,d\xi.\end{aligned}$$

6.2 Filtering mechanisms for the quadratic finite element method

Set $\Omega=\mathbb{R}\setminus[-1,1]$. In the following, we will design classes of initial data in which the semi-discrete observability inequality

$$E_h(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1})\le C_h(T)\int\limits_0^T E_{\Omega,h}(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1},t)\,dt \qquad (6.16)$$

holds uniformly as $h\to 0$.

The first two filtering algorithms deal with initial data in (6.2) $\overrightarrow{U}^{h,i}$, $i=0,1$, concentrated on the acoustical components, i.e.

$$\widehat{U}^{h,i}(\xi)=P_h^a(\xi)\widehat{u}^{h,i}(\xi),\ i=0,1,\xi\in\Pi_h. \qquad (6.17)$$

In this case, the solution of (6.2) has a simpler form with respect to the one indicated in (6.14), i.e.

$$\widehat{U}^h(\xi,t)=P_h^a(\xi)\sum_{\pm}\frac{1}{2}\Big(\widehat{u}^{h,0}(\xi)\pm\frac{\widehat{u}^{h,1}(\xi)}{i\lambda_h^a(\xi)}\Big)\exp(\pm it\lambda_h^a(\xi)), \qquad (6.18)$$

whose total energy is

$$E_h(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1})=\frac{1}{4\pi}\int\limits_{\Pi_h}(m_h^a(\xi)|\widehat{u}^{h,1}(\xi)|^2+r_h^a(\xi)|\widehat{u}^{h,0}(\xi)|^2)\,d\xi, \qquad (6.19)$$

where

$$\begin{aligned}m_h^a(\xi)&=\langle M_h(\xi)P_h^a(\xi),P_h^a(\xi)\rangle\\
&=\frac{8}{15}\frac{\Big(1+9\cos^2\Big(\frac{\xi h}{2}\Big)+\sqrt{\Delta_h(\xi)}\Big)^2+125\cos^2\Big(\frac{\xi h}{2}\Big)\Big(1+\sin^2\Big(\frac{\xi h}{2}\Big)\Big)}{400\cos^2\Big(\frac{\xi h}{2}\Big)+\Big(1+4\cos^2\Big(\frac{\xi h}{2}\Big)+\sqrt{\Delta_h(\xi)}\Big)^2}\end{aligned}$$

and

$$r_h^a(\xi) = \langle R_h(\xi)P_h^a(\xi), P_h^a(\xi)\rangle = \frac{16}{3h^2}\frac{\left(1-16\cos^2\left(\frac{\xi h}{2}\right)+\sqrt{\Delta_h(\xi)}\right)^2+300\sin^2\left(\frac{\xi h}{2}\right)\cos^2\left(\frac{\xi h}{2}\right)}{400\cos^2\left(\frac{\xi h}{2}\right)+\left(1+4\cos^2\left(\frac{\xi h}{2}\right)+\sqrt{\Delta_h(\xi)}\right)^2}.$$

Lemma 6.2.1. *For all $\xi \in \Pi_h$, $r_h^a(\xi) = \Lambda_h^a(\xi)m_h^a(\xi)$.*

The proof of Lemma 6.2.1 follows the one of Lemma 4.4.1 in Chapter 2.

For $\delta \in (0,1)$, set $\Pi_h^\delta := [-\pi\delta/h, \pi\delta/h]$. Let us define the space of Fourier filtered data:

$$I_h^\delta = \{\overrightarrow{f} \in \ell^2(h\mathbb{Z}) : \mathrm{supp}(\widehat{f^h}) \subset \Pi_h^\delta\}. \tag{6.20}$$

For $\overrightarrow{f} \in \ell^2(h\mathbb{Z})$, define the projection on I_h^δ of $\overrightarrow{f}$ to be

$$\Gamma_h^\delta f_j = \frac{1}{2\pi}\int_{\Pi_h^\delta} \widehat{f^h}(\xi)\exp(i\xi x_j)\,d\xi. \tag{6.21}$$

We say that the initial data $\overrightarrow{u}^{h,i}$, $i = 0, 1$, in (6.17) is filtered with parameter $\delta \in (0,1)$ if

$$\overrightarrow{u}^{h,i} \in I_h^\delta, \quad \forall i = 0, 1. \tag{6.22}$$

Since $\lambda_h^a(\xi)$ is increasing in ξ, if the initial data in (6.2) is such that $\overrightarrow{U}^{h,i} \in (I_h^\delta)^2$, $i = 0, 1$, and such that (6.17) holds, then the maximum frequency involved in the solution (6.18) is $\lambda_h^a(\delta\pi/h)$.

Another kind of initial data $\overrightarrow{u}^{h,i}$ used in this section is obtained by linear interpolation from a coarse grid of size $2h$, G^{2h}, to a fine grid of size h, G^h. Explicitly, the odd components are related to the even ones in the following way:

$$u_{2j}^{h,i} = \frac{u_{2j-1}^{h,i} + u_{2j+1}^{h,i}}{2}, i = 0, 1. \tag{6.23}$$

The following two results of uniform observability hold:

Theorem 6.2.1 (Concentration on the acoustical mode + Fourier filtering)**.** *In (6.2) consider initial data $\overrightarrow{U}^{h,i}$, $i = 0, 1$, concentrated on the acoustical mode, i.e. verifying condition (6.17), and filtered with parameter $\delta \in (0,1)$, i.e. verifying (6.22). If $T > T^{a,\delta}$, with*

$$T^{a,\delta} = \frac{2}{\min\limits_{\xi\in\Pi_h^\delta}\partial_\xi\lambda_h^a(\xi)}(1 + C^{a,\delta})$$

and $C^{a,\delta} \in (0,1)$, then the observability inequality (6.16) holds uniformly as $h \to 0$, with an observability constant

$$C_h(T) = C^{a,\delta}(T) = \frac{1}{T - T^{a,\delta}}.$$

Theorem 6.2.2 (Concentration on the acoustical mode + bi-grid algorithm)**.** *In (6.2), consider initial data $\overrightarrow{U}^{h,i}$, $i = 0, 1$, concentrated on the acoustical mode, i.e. verifying the condition (6.17), such that $\overrightarrow{u}^{h,i}$, $i = 0, 1$, verify the bi-grid condition (6.23). If $T > T^{a,1/2}$, with $T^{a,\delta}$ given in Theorem 6.2.1, then the observability inequality (6.16) holds uniformly as $h \to 0$.*

Proof of Theorem 6.2.1. The proof of this theorem is based on the same arguments in the one of Theorem 4.5.1 in Chapter 2. However, we expose here the main idea. Firstly, for $I = [-1, 1]$ we have the identity

$$\int_0^T E_{\Omega,h}(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\,dt = TE_h(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}) - \int_0^T E_{I,h}(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\,dt.$$

Also

$$\int_0^T E_{I,h}(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\, dt = \sum_{k=1}^{8} \int_0^T E_{I,h}^k(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\, dt \leq \sum_{k=1}^{8} I_k,$$

with

$$I_k = \int_{\mathbb{R}} E_{I,h}^k(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}, t)\, dt.$$

By the change of variable $\zeta = \lambda_h^a(\xi)$, which is well-defined taking into account the fact that $\lambda_h^a(\xi)$ is injective on Π_h, we transform the solution (6.18) of (6.4) into an inverse Fourier transform in time as follows

$$N_j(t) = \frac{1}{2\pi} \int_{\lambda_h^a(\Pi_h^\delta)} \frac{p_{h,1}^a(\xi(\zeta))}{i\lambda_h^a(\xi(\zeta))} \widehat{U}_j^{h,-}(\xi(\zeta)) \exp(it\zeta)\, d\zeta$$

and

$$H_{j+1/2}(t) = \frac{1}{2\pi} \int_{\lambda_h^a(\Pi_h^\delta)} \frac{p_{h,2}^a(\xi(\zeta))}{i\lambda_h^a(\xi(\zeta))} \widehat{U}_{j+1/2}^{h,-}(\xi(\zeta)) \exp(it\zeta)\, d\zeta,$$

with

$$d\zeta = \frac{d\xi}{\partial_\xi \lambda_h^a(\xi(\zeta))},$$

$\xi(\zeta) = (\lambda_h^a)^{-1}(\zeta)$,

$$\widehat{u}^{h,\pm}(\xi) = \frac{1}{2}\big(i\lambda_h^a(\xi)\widehat{u}^{h,0}(\xi) \pm \widehat{u}^{h,1}(\xi)\big)$$

and

$$\widehat{U}_\alpha^{h,+}(\xi) = \sum_{\pm} \widehat{u}^{h,\pm}(\pm\xi) \exp(\pm i\xi x_\alpha), \quad \widehat{U}_\alpha^{h,-}(\xi) = \sum_{\pm} [\pm \widehat{u}^{h,\pm}(\pm\xi) \exp(\pm i\xi x_\alpha)].$$

By applying the Parseval identity in time for each quadratic term in each I_k, $k = 1, \cdots, 8$, and then undoing the change of variable, we conclude that

$$\sum_{k=1}^{8} I_k = h \sum_{x_j \in I} \frac{1}{2\pi} \int_{\Pi_h^\delta} \big[\alpha_h^+(\xi)|\widehat{U}_j^{h,+}(\xi)|^2 + \alpha_h^-(\xi)|\widehat{U}_j^{h,-}(\xi)|^2\big] \frac{1}{\partial_\xi \lambda_h^a(\xi)}\, d\xi,$$

with

$$\alpha_h^+(\xi) = \sin^2\Big(\frac{\xi h}{2}\Big) m_h^a(\xi) + \sin^2\Big(\frac{\xi h}{2}\Big)\Big(\frac{2\cos(\xi h)}{15} + \frac{2\cos(\xi h)}{h^2\Lambda_h^a(\xi)} - \frac{1}{30}\Big)|p_{h,1}^a(\xi)|^2$$

and

$$\alpha_h^-(\xi) = \cos^2\Big(\frac{\xi h}{2}\Big) m_h^a(\xi) - \sin^2\Big(\frac{\xi h}{2}\Big)\Big(\frac{2\cos(\xi h)}{15} + \frac{2\cos(\xi h)}{h^2\Lambda_h^a(\xi)} - \frac{1}{30}\Big)|p_{h,1}^a(\xi)|^2.$$

Observe that $\alpha_h^\pm(\xi) \geq 0$, $\alpha_h^+(\xi) + \alpha_h^-(\xi) = m_h^a(\xi)$ and $|\alpha_h^+(\xi) - \alpha_h^-(\xi)| = \sin^2(\xi h/2)\alpha_h(\xi)$, with

$$\alpha_h(\xi) = \frac{40\left|2\cos^2\left(\frac{\xi h}{2}\right)\frac{4\cos(\xi h)-1}{3} - \frac{\cos(\xi h)\left(1+9\cos^2\left(\frac{\xi h}{2}\right)+\sqrt{\Delta_h(\xi)}\right)\left(1+\sqrt{\Delta_h(\xi)}\right)}{71-11\cos^2\left(\frac{\xi h}{2}\right)+4\sqrt{\Delta_h(\xi)}}\right|}{400\cos^2\left(\frac{\xi h}{2}\right) + \left(1 + 4\cos^2\left(\frac{\xi h}{2}\right) + \sqrt{\Delta_h(\xi)}\right)^2}.$$

Since

$$|\widehat{U}_j^{h,+}(\xi)|^2 + |\widehat{U}_j^{h,-}(\xi)|^2 = 2(|\widehat{u}^{h,+}(\xi)|^2 + |\widehat{u}^{h,-}(-\xi)|^2),$$

for all $j \in \mathbb{Z}$ and all $\xi \in \Pi_h$, $m_h^a(\xi)$, $\alpha_h(\xi)$, $|\sin(\xi h/2)|$ and $\partial_\xi \lambda_h^a(\xi)$ are even functions on Π_h^δ, the fact that

$$|\widehat{u}^{h,+}(\xi)|^2 + |\widehat{u}^{h,-}(\xi)|^2 = \frac{1}{2}[\Lambda_h^a(\xi)|\widehat{u}^{h,0}(\xi)|^2 + |\widehat{u}^{h,1}(\xi)|^2]$$

and Lemma 6.2.1, we get

$$h\sum_{x_j\in I}\frac{1}{2\pi}\int\limits_{\Pi_h^\delta}\left[\alpha_h^+(\xi)|\widehat{U}_j^{h,+}(\xi)|^2+\alpha_h^-(\xi)|\widehat{U}_j^{h,-}(\xi)|^2\right]\frac{1}{\partial_\xi\lambda_h^a(\xi)}\,d\xi$$
$$=\frac{1}{2\pi}\int\limits_{\Pi_h^\delta}\left(m_h^a(\xi)|\widehat{u}^{h,1}(\xi)|^2+r_h^a(\xi)|\widehat{u}^{h,0}(\xi)|^2\right)\frac{1+C_h^a(\xi)}{\partial_\xi\lambda_h^a(\xi)}\,d\xi,$$

with

$$C_h^a(\xi)=\sin^2\Big(\frac{\xi h}{2}\Big)\frac{\alpha_h(\xi)}{m_h^a(\xi)}.$$

We conclude by considering $C^{a,\delta}=\|C_h^a\|_{L^\infty(\Pi_h^\delta)}=\|C_1^a\|_{L^\infty(\Pi_1^\delta)}$, which does not depend on h. Also, by a rescaling argument, $\min\limits_{\xi\in\Pi_h^\delta}\partial_\xi\lambda_h^a(\xi)=\min\limits_{\xi\in\Pi_1^\delta}\partial_\xi\lambda_1^a(\xi)$ does not depend on h. Observe that $C_h^a(\xi)\le(\alpha_h^+(\xi)+\alpha_h^-(\xi))/m_h^a(\xi)=1$ and that $C^{a,\delta}\to 0$ as $\delta\to 0$ like $\sin^2(\delta\pi/2)$. □

The proof of Theorem 6.2.2 is based on the same arguments as the ones in the proof of Theorem 4.5.2, in Chapter 2, i.e. a dyadic decomposition argument. We only have to prove the analogous of Proposition 5.1, which is as follows:

Proposition 6.2.1. *In (6.2), consider initial data $\overrightarrow{U}^{h,i}$, $i=0,1$, satisfying the condition (6.17) and s.t. $\overrightarrow{u}^{h,i}$, $i=0,1$, verify the bi-grid condition (6.23). Then*

$$E_h(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1})\le 2E_h(\Gamma_h^{1/2}\overrightarrow{U}^{h,0},\Gamma_h^{1/2}\overrightarrow{U}^{h,1}).\tag{6.24}$$

Proof of Proposition 6.2.1. From the bi-grid condition (6.23), we get

$$\widehat{u}^{h,i}(\xi)=\cos^2\Big(\frac{\xi h}{2}\Big)\widehat{u}^{2h,i}(\xi),\forall\xi\in\Pi_h,\forall i=0,1,$$

where $\widehat{u}^{2h,i}(\xi)$ is a π/h-periodic function, the SDFT at scale $2h$ of the sequence $\overrightarrow{u}^{2h,i}=(u_{2j+1}^i)_{j\in\mathbb{Z}}$. Then, by using the fact that $m_h^a(\xi)$ and $r_h^a(\xi)$ are even functions, we decompose the total energy as follows:

$$E_h(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1})=E_h(\Gamma_h^{1/2}\overrightarrow{U}^{h,0},\Gamma_h^{1/2}\overrightarrow{U}^{h,1})+\frac{1}{4\pi}\int\limits_{\Pi_{2h}}\frac{m_h^a(\xi+\pi/h)}{m_h^a(\xi)}\tan^4\Big(\frac{\xi h}{2}\Big)m_h^a(\xi)|\widehat{u}^{h,1}(\xi)|^2\,d\xi$$
$$+\frac{1}{4\pi}\int\limits_{\Pi_{2h}}\frac{r_h^a(\xi+\pi/h)}{r_h^a(\xi)}\tan^4\Big(\frac{\xi h}{2}\Big)r_h^a(\xi)|\widehat{u}^{h,0}(\xi)|^2\,d\xi.$$

From the expression of the total energy (6.19), we see that in order to conclude the proof of this proposition, is sufficient to show that:

$$\Big|\frac{m_h^a(\xi+\pi/h)}{m_h^a(\xi)}\tan^4\Big(\frac{\xi h}{2}\Big)\Big|\le 1\text{ and }\Big|\frac{r_h^a(\xi+\pi/h)}{r_h^a(\xi)}\tan^4\Big(\frac{\xi h}{2}\Big)\Big|\le 1.$$

Consider the function $f:(0,1)\to\mathbb{R}$ defined by $f(x)=f_1(x)f_2(x)$, where

$$f_1(x)=\frac{x^2}{400x+(1+4x+\sqrt{1+268x-44x^2})^2},\ f_2(x)=125x(2-x)+(1+9x+\sqrt{1+268x-44x^2})^2.$$

It is very easy to see that

$$\frac{m_h^a(\xi+\pi/h)}{m_h^a(\xi)}\tan^4\Big(\frac{\xi h}{2}\Big)=\frac{f(x(\xi))}{f(1-x(\xi))},\quad x(\xi)=\sin^2\Big(\frac{\xi h}{2}\Big),$$

and that both $f_1(x)$ and $f_2(x)$ are increasing functions on $(0,1)$, so f is also an increasing function on $(0,1)$. Then, since $x(\xi)$ is increasing on Π_{2h}, the same could be said about $f(x(\xi))$ and $1/f(1-x(\xi))$ and therefore on $f(x(\xi))/f(1-x(\xi))$, the maximum of the last function being attained at $\xi=\pi/2h$ and equal to 1.

On the other hand,

$$\frac{r_h^a(\xi+\pi/h)}{r_h^a(\xi)}\tan^4\Big(\frac{\xi h}{2}\Big)=\frac{g(x(\xi))}{g(1-x(\xi))},$$

with

$$g(x)=\frac{300x(1-x)+(1-16x+\sqrt{1+268x-44x^2})^2}{400x+(1+4x+\sqrt{1+268x-44x^2})^2}\frac{x}{1-x}.$$

It is not difficult to see that

$$g'(x)(1-x)^2\big(400x+(1+4x+\sqrt{1+268x-44x^2})^2\big)^2=\frac{g_1(x)\sqrt{1+268x-44x^2}+g_2(x)}{\sqrt{1+268x-44x^2}},$$

where $g_1(x)$ is a forth-order polynomial strictly positive on $(0,1)$ and $g_2(x)$ is a fifth-order polynomial such that $g_2(0)>0$ and $g_2(1)<0$.

For $x\in(0,1)$ such that $g_2(x)\geq 0$, it is clear that $g'(x)>0$ and g increasing.

For $x\in(0,1)$ such that $g_2(x)<0$, it can be proved that

$$g_1^2(x)(1+268x-44x^2)-g_2^2(x)=x^2(1-x)^2(x-\alpha)(\beta-x)g_3(x)g_4(x),$$

where $g_3(x)$, $g_4(x)$ are positive second-order polynomials, $\alpha<0$, $\beta>4$, then $g_1^2(x)(1+268x-44x^2)-g_2^2(x)>0$ on $(0,1)$. This means that g is an increasing function on $(0,1)$ and then $g(x(\xi))/g(1-x(\xi))$ is increasing, its maximum value on $\xi\in\Pi_{2h}$ being attained at $\xi=\pi/2h$ and equal to 1. This concludes the proof of Proposition 6.2.1. □

In the following two filtering mechanisms, the initial data in (6.2), $\overrightarrow{U}^{h,i}=(N_j^i,H_{j+1/2}^i)_{j\in\mathbb{Z}}$, is a linear one, i.e.

$$H_{j+1/2}^i=\frac{N_{j+1}^i+N_j^i}{2}\ \text{ or }\ \widehat{H}^{h,i}(\xi)=\cos\Big(\frac{\xi h}{2}\Big)\widehat{N}^{h,i}(\xi),\forall\xi\in\Pi_h,i=0,1. \tag{6.25}$$

Replacing (6.25) into (6.15), the total energy corresponding to this choice of the initial data coincides with the total energy corresponding to the classical FEM method:

$$E_h(\overrightarrow{U}^{h,0},\overrightarrow{U}^{h,1})=\frac{1}{4\pi}\int\limits_{\Pi_h}\Big[\frac{2+\cos(\xi h)}{3}|\widehat{N}^{h,1}(\xi)|^2+\frac{4}{h^2}\sin^2\Big(\frac{\xi h}{2}\Big)|\widehat{N}^{h,0}(\xi)|^2\Big]\,d\xi. \tag{6.26}$$

Also the solution of (6.2) simplifies with respect to the one indicated in (6.14) as follows:

$$\overrightarrow{U}^h(\xi,t)=\frac{1}{\det(P_h(\xi))}P_h^{lin}(\xi)\begin{pmatrix}\widehat{N}_a^h(\xi,t)\\ \widehat{N}_o^h(\xi,t)\end{pmatrix}, \tag{6.27}$$

where $\det(P_h(\xi))=p_{h,1}^a(\xi)p_{h,2}^o(\xi)-p_{h,1}^o(\xi)p_{h,2}^a(\xi)$, $P_h(\xi)$ is the matrix defined in (6.13),

$$P_h^{lin}(\xi)=\begin{pmatrix}p_{h,1}^a(\xi)\Big(p_{h,2}^o(\xi)-p_{h,1}^o(\xi)\cos\Big(\frac{\xi h}{2}\Big)\Big) & p_{h,1}^o(\xi)\Big(-p_{h,2}^a(\xi)+p_{h,1}^a(\xi)\cos\Big(\frac{\xi h}{2}\Big)\Big)\\ p_{h,2}^a(\xi)\Big(p_{h,2}^o(\xi)-p_{h,1}^o(\xi)\cos\Big(\frac{\xi h}{2}\Big)\Big) & p_{h,2}^o(\xi)\Big(-p_{h,2}^a(\xi)+p_{h,1}^a(\xi)\cos\Big(\frac{\xi h}{2}\Big)\Big)\end{pmatrix},$$

$$\widehat{N}_a^h(\xi,t)=\frac{1}{2}\sum_{\pm}\Big[\widehat{N}^{h,0}(\xi)\pm\frac{\widehat{N}^{h,1}(\xi)}{i\lambda_h^a(\xi)}\Big]\exp(\pm it\lambda_h^a(\xi)),$$

$$\widehat{N}_o^h(\xi,t)=\frac{1}{2}\sum_{\pm}\Big[\widehat{N}^{h,0}(\xi)\pm\frac{\widehat{N}^{h,1}(\xi)}{i\lambda_h^o(\xi)}\Big]\exp(\pm it\lambda_h^o(\xi)).$$

Before stating the following two results of this section, let us introduce some useful notions. If $\overrightarrow{f}^h(t)$ is an evolution process such that its SDFT,

$$\widehat{f}^h(\xi,t)=\widehat{f}_a^h(\xi)\exp(it\lambda_h^a(\xi))+\widehat{f}_o^h(\xi)\exp(it\lambda_h^o(\xi)),\ \xi\in\Pi_h, \tag{6.28}$$

involves the two branches of frequencies $\lambda_h^a(\xi)$ and $\lambda_h^o(\xi)$ and $\widehat{f}_a^h(\xi)$ and $\widehat{f}_o^h(\xi)$ are scalar functions, we define its projection on the acoustical branch to be

$$\Gamma_a f_j(t)=\frac{1}{2\pi}\int_{\Pi_h}\widehat{f}_a^h(\xi)\exp(it\lambda_h^a(\xi))\exp(i\xi x_j)\,d\xi. \tag{6.29}$$

If $\overrightarrow{f}^h(t)$ is a vectorial process, in the sense that both $\widehat{f}_a^h(\xi)$ and $\widehat{f}_o^h(\xi)$ are vectorial functions with the same number of components, we define $\Gamma_a\overrightarrow{f}^h(t)$ to be the vector obtained as Γ_a acting on each one of the components of $\overrightarrow{f}^h(t)$.

On the vectorial process $\overrightarrow{U}^h(t)$ given by (6.27), the projection Γ_a acts as follows

$$\Gamma_a U_j(t)=\frac{1}{2\pi}\int_{\Pi_h}P_h^a(\xi)\frac{p_{h,2}^o(\xi)-p_{h,1}^o\cos\left(\frac{\xi h}{2}\right)}{p_{h,1}^a(\xi)p_{h,2}^o(\xi)-p_{h,1}^o(\xi)p_{h,2}^a(\xi)}\widehat{N}_a^h(\xi,t)\exp(i\xi x_j)\,d\xi, \tag{6.30}$$

where $P_h^a(\xi)$ is the eigenvector corresponding to the acoustical eigenvalue $\Lambda_h^a(\xi)$.

Note that $\Gamma_a\overrightarrow{U}^h(t)$ verifies the wave equation (6.2) with the initial data $\overrightarrow{V}^{h,i}$, $i=0,1$, such that

$$\widehat{V}^{h,i}(\xi)=P_h^a(\xi)\frac{p_{h,2}^o(\xi)-p_{h,1}^o\cos\left(\frac{\xi h}{2}\right)}{p_{h,1}^a(\xi)p_{h,2}^o(\xi)-p_{h,1}^o(\xi)p_{h,2}^a(\xi)}\widehat{N}^{h,i}(\xi). \tag{6.31}$$

Then the total energy of $\Gamma_a\overrightarrow{U}^h(t)$, $E_h(\Gamma_a\overrightarrow{U}^h(t),\partial_t\Gamma_a\overrightarrow{U}^h(t))$, is conserved in time. We denote it by $E_h(\Gamma_a\overrightarrow{U}^{h,0},\Gamma_a\overrightarrow{U}^{h,1})$. Observe that $\widehat{V}^{h,i}(\xi)$, $i=0,1$, verifies the condition (6.17), with

$$\widehat{u}^{h,i}(\xi)=\frac{p_{h,2}^o(\xi)-p_{h,1}^o\cos\left(\frac{\xi h}{2}\right)}{p_{h,1}^a(\xi)p_{h,2}^o(\xi)-p_{h,1}^o(\xi)p_{h,2}^a(\xi)}\widehat{N}^{h,i}(\xi),\ i=0,1.$$

Then the total energy $E_h(\Gamma_a\overrightarrow{U}^{h,0},\Gamma_a\overrightarrow{U}^{h,1})$ can be computed using (6.19) and is given by

$$\begin{aligned}&E_h(\Gamma_a\overrightarrow{U}^{h,0},\Gamma_a\overrightarrow{U}^{h,1})= \\ &=\frac{1}{4\pi}\int_{\Pi_h}\Big|\frac{p_{h,2}^o(\xi)-p_{h,1}^o\cos\left(\frac{\xi h}{2}\right)}{p_{h,1}^a(\xi)p_{h,2}^o(\xi)-p_{h,1}^o(\xi)p_{h,2}^a(\xi)}\Big|^2\left(m_h^a(\xi)|\widehat{N}^{h,1}(\xi)|^2+r_h^a(\xi)|\widehat{N}^{h,0}(\xi)|^2\right)d\xi.\end{aligned} \tag{6.32}$$

The following two results of uniform observability hold:

Theorem 6.2.3 (Linear initial data + nodal part given by Fourier filtering)**.** *In (6.2), consider linear initial data $\overrightarrow{U}^{h,i}=(\overrightarrow{N}^{h,i},\overrightarrow{H}^{h,i})'$, $i=0,1$, i.e. $\overrightarrow{N}^{h,i}$ and $\overrightarrow{H}^{h,i}$ are related by (6.25) and s.t. $\overrightarrow{N}^{h,i}\in I_h^\delta$, $\delta\in(0,1)$. Then for all $T>T^{a,\delta}$, with $T^{a,\delta}$ introduced in Theorem 6.2.1, the observability inequality (6.16) holds uniformly as $h\to 0$.*

In the following theorem, the nodal initial data is given by a bi-grid algorithm, i.e.

$$N_{2j}^i=\frac{N_{2j-1}^i+N_{2j+1}^i}{2},i=0,1. \tag{6.33}$$

Theorem 6.2.4 (Linear initial data + nodal part given by a bi-grid algorithm)**.** *In (6.2), consider linear initial data* $\overrightarrow{U}^{h,i} = (\overrightarrow{N}^{h,i}, \overrightarrow{H}^{h,i})'$, $i = 0, 1$, *i.e.* $\overrightarrow{N}^{h,i}$ *and* $\overrightarrow{H}^{h,i}$ *related by (6.25) and s.t.* $\overrightarrow{N}^{h,i}$ *verify (6.33). For all* $T > T^{a,1/2}$, *with* $T^{a,\delta}$ *given in Theorem 6.2.1, the observability inequality (6.16) holds uniformly as* $h \to 0$.

The proof of Theorem 6.2.3 follows a dyadic decomposition argument described by the same steps as the one of Theorem 4.5.3 in Chapter 2, but Proposition 4.5.2 has to be replaced by the following one:

Proposition 6.2.2. *In (6.2), consider linear initial data* $\overrightarrow{U}^{h,i} = (\overrightarrow{N}^{h,i}, \overrightarrow{H}^{h,i})'$, $i = 0, 1$, *i.e.* $\overrightarrow{N}^{h,i}$ *and* $\overrightarrow{H}^{h,i}$ *are related by (6.25) and s.t.* $\overrightarrow{N}^{h,i} \in I_h^\delta$, $\delta \in (0, 1)$. *Then there exists a constant* $C(\delta) > 0$ *independent of* h *s.t.*

$$E_h(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}) \leq C(\delta) E_h(\Gamma_a \overrightarrow{U}^{h,0}, \Gamma_a \overrightarrow{U}^{h,1}). \tag{6.34}$$

Moreover, $C(\delta) \to \infty$ *as* $\delta \to 1$.

Proof of Proposition 6.2.2. From the explicit expressions of the total energy (6.26) and of the energy projected on the acoustical component (6.32), we observe that it is sufficient to prove that both

$$m_h(\xi) = \frac{\frac{2+\cos(\xi h)}{3}}{\left|\frac{p^o_{h,2}(\xi) - p^o_{h,1}\cos\left(\frac{\xi h}{2}\right)}{p^a_{h,1}(\xi)p^o_{h,2}(\xi) - p^o_{h,1}(\xi)p^a_{h,2}(\xi)}\right|^2 m^a_h(\xi)} \text{ and } r_h(\xi) = \frac{\frac{4}{h^2}\sin^2\left(\frac{\xi h}{2}\right)}{\left|\frac{p^o_{h,2}(\xi) - p^o_{h,1}\cos\left(\frac{\xi h}{2}\right)}{p^a_{h,1}(\xi)p^o_{h,2}(\xi) - p^o_{h,1}(\xi)p^a_{h,2}(\xi)}\right|^2 r^a_h(\xi)}$$

are bounded on I_h^δ. It is easy to show that $m_h(\xi)$ and $r_h(\xi)$ can be written in an explicit way as follows:

$$m_h(\xi) = \frac{A\Big(\cos^2\left(\frac{\xi h}{2}\right)\Big)}{\cos^2\left(\frac{\xi h}{2}\right)} \text{ and } r_h(\xi) = \frac{B\Big(\cos^2\left(\frac{\xi h}{2}\right)\Big)}{\cos^2\left(\frac{\xi h}{2}\right)},$$

with

$$A(x) = \frac{5}{8}\frac{1+2x}{(14+x+\sqrt{1+268x-44x^2})^2}\frac{[(1+4x+\sqrt{1+268x-44x^2})^2 + 20x(13-3x)]^2}{(1+9x+\sqrt{1+268x-44x^2})^2 + 125x(2-x)}$$

and

$$B(x) = \frac{1}{100}\frac{[(1+4x+\sqrt{1+268x-44x^2})^2 + 20x(13-3x)]^2}{(14+x+\sqrt{1+268x-44x^2})^2\left[4x + \frac{75(1-x)(1+\sqrt{1+268x-44x^2})^2}{(71-11x+\sqrt{1+268x-44x^2})^2}\right]}.$$

Observe that by the change of variable $\eta = h\xi \in \Pi_1^\delta$, $m_h(\xi) = m_1(\eta)$ and $r_h(\xi) = r_1(\eta)$ and then

$$\|m_h\|_{L^\infty(\Pi_h^\delta)} = \|m_1\|_{L^\infty(\Pi_1^\delta)} \text{ and } \|r_h\|_{L^\infty(\Pi_h^\delta)} = \|r_1\|_{L^\infty(\Pi_1^\delta)}$$

do not depend on h.

Also $A, B \in L^\infty(0, 1)$ and then

$$\|m_h\|_{L^\infty(\Pi_h^\delta)} \leq \frac{\|A\|_{L^\infty(0,1)}}{\cos^2(\pi\delta/2)} \text{ and } \|r_h\|_{L^\infty(\Pi_h^\delta)} \leq \frac{\|B\|_{L^\infty(0,1)}}{\cos^2(\pi\delta/2)}.$$

We conclude the proof of Proposition 6.2.2 by considering

$$C(\delta) = \frac{\max\{\|A\|_{L^\infty(0,1)}, \|B\|_{L^\infty(0,1)}\}}{\cos^2(\pi\delta/2)}.$$

□

The proof of Theorem 6.2.4 follows the same methodology as the one of Theorem 4.5.4 in Chapter 2. Proposition 4.5.3 in Step 1 has to be replaced by the following one:

Proposition 6.2.3. *In (6.2), consider linear initial data* $\overrightarrow{U}^{h,i} = (\overrightarrow{N}^{h,i}, \overrightarrow{H}^{h,i})'$, $i = 0, 1$, *i.e.* $\overrightarrow{N}^{h,i}$ *and* $\overrightarrow{H}^{h,i}$ *related by (6.25) and s.t.* $\overrightarrow{N}^{h,i}$ *verify (6.33). Then*

$$E_h(\overrightarrow{U}^{h,0}, \overrightarrow{U}^{h,1}) \leq 2E_h(\Gamma_h^{1/2}\overrightarrow{U}^{h,0}, \Gamma_h^{1/2}\overrightarrow{U}^{h,1}). \tag{6.35}$$

Also Proposition 6.2.2 with $\delta = 1/2$ has to be applied in Step 2.

Chapter 7

Comentarios y problemas abiertos

En esta tesis hemos analizado los siguiente problemas:

1. Construcción de soluciones numéricas en altas frecuencias y sus consecuencias sobre el orden de divergencia de la constante de observabilidad para las aproximaciones mediante diferencias finitas de la ecuación de ondas.

2. Análisis de Fourier y mecanismos de filtrado para las semi-discretizaciones utilizando dos clases de métodos de Galerkin discontinuos (GD) (SIPG y LDG) de la ecuación de ondas $1-d$.

3. Análisis de Fourier de las semi-discretizaciones con funciones spline de ecuación de ondas $1-d$.

4. Análisis de Fourier y mecanismos de filtrado para la semi-discretización mediante elementos finitos cuadráticos clásicos de la ecuación de ondas $1-d$.

Para las aproximaciones mediante diferencias finitas de la ecuación de ondas, hemos construido soluciones para las cuales la constante de observabilidad explota de manera polinomial arbitraria con respecto al paso del mallado.

Para las semi-discretizaciones con funciones de clase C^k, hemos mostrado que la correspondiente velocidad de grupo se anula cuando $\xi = \pm\pi/h$, lo que permite construir soluciones que se propagan a velocidad arbitrariamente pequeña. Usando el truncamiento en Fourier como algoritmo de filtrado, se puede mostrar que a medida que k aumenta, el tiempo de observabilidad mejora con respeto al obtenido con elementos finitos clásicos lineales.

Para las semi-discretizaciones mediante métodos de GD o elementos finitos clásicos cuadráticos, en vista del análisis de Fourier, mostramos que hay paquetes de ondas que se propagan a velocidad arbitrariamente pequeña, tanto en el diagrama de dispersión físico, como en el espúreo y concluimos que la correspondiente constante de observabilidad explota al menos de manera polinomial arbitraria. Tambien diseñamos clases de datos iniciales obtenidos por truncamientos en Fourier o algoritmos bimalla, para las cuales las propiedades de observabilidad se recuperan de manera uniforme con respeto al paso del mallado. Estos mecanismos de filtrado tienen dos objetivos: proyectar la solución numérica en el diagrama físico y usar mecanismos de filtrados para eliminar el comportamiento a altas frecuencias en este diagrama.

En lo que sigue, presentamos de manera sistemática algunos de los problemas abiertos relacionados con los temas que hemos tratado a lo largo de la tesis.

- **La divergencia exponencial de la constante de observabilidad.** Como hemos probado en la Proposición 3.2.1 en el primer capítulo, a través de un resultado de [42] que se refiere al problema de observabilidad para la ecuación de ondas semi-discreta en un intervalo de longitud finita, se puede deducir que en la version $1-d$ de (1.1) existen datos iniciales para los cuales la constante de observabilidad $C_h(T)$ definida por (1.3) explota de manera exponencial con respeto al paso del mallado. Pero dicho resultado no ofrece ninguna información sobre la

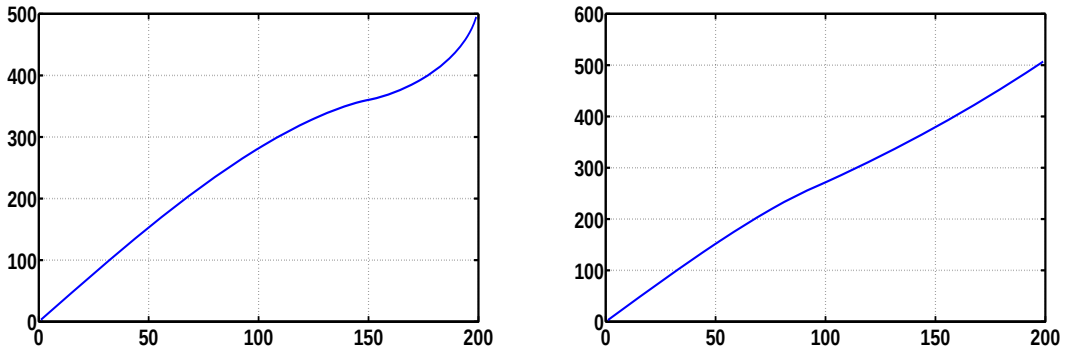

Figure 7.1: Discrete eigenvalues on two particular non-uniform meshes.

forma de la correspondiente solución y no se puede generalizar a varias dimensiones espaciales o condiciones de borde de otra naturaleza. Nuestras construcciones de paquetes de onda en el primer capítulo conducen a ordenes de divergencia polinomiales para $C_h(T)$. Sería interesante desarrollar aún más estas construcciones y las correspondientes estimaciones para verificar si la constante de observabilidad $C_h(T)$ explota de manera exponencial en situaciones más generales.

- **Propiedades de propagación de las aproximaciones de la ecuación de ondas en medios heterogéneos discretos.** En general, el análisis de las propiedades de propagación y dispersión de las aproximaciones de la ecuación de ondas está basado en el análisis de Fourier de la solución numérica. Para esquemas numéricos con coeficientes constantes en mallados uniformes, este asunto está entendido bastante bien. Se sabe poco sobre esto en medios heterogéneos discretos representados por mallas irregulares o coeficientes variables, incluso en el caso de los esquemas más clásicos como las diferencias finitas o el método de los elementos finitos clásicos lineales. Todos estos asuntos se constituyen en temas difíciles en los cuales incluso problemas elementales como la definición del símbolo del laplaciano discreto están todavía abiertos, excepto algunos casos particulares obtenidos por deformaciones regulares de mallas uniformes o periódicas ([10], [23]). Por consiguiente, la adaptación de nuestras construcciones de paquetes de ondas altamente oscilantes en este contexto más general constituye un problema abierto muy estimulante.

  Otro tema interesante consiste en investigar la posibilidad de construir mallas adaptadas, en las cuales las propiedades de propagación de las ondas disretas se mantienen de manera uniforme con respeto al paso del mallado. Consideramos una malla uniforme $x_j = jh$ de $(0,1)$, $0 \leq j \leq N+1$ y una no-uniforme y_j obtenida como $y_j = f(x_j)$, con $f : (0,1) \to (0,1)$. En la malla no-uniforme, consideramos el problema semi-discreto:

$$\begin{cases} \partial_{tt}\phi_j(t) - \left(\frac{2(\phi_{j+1}(t)-\phi_j(t))}{(y_{j+1}-y_j)(y_{j+1}-y_{j-1})} - \frac{2(\phi_j(t)-\phi_{j-1}(t))}{(y_j-y_{j-1})(y_{j+1}-y_{j-1})}\right) = 0, \quad 1 \leq j \leq N \\ \phi_0 = \phi_{N+1} = 0, \\ \phi_j(0) = \phi_j^0, \quad \partial_t\phi_j(0) = \phi_j^1, \qquad\qquad 1 \leq j \leq N. \end{cases} \tag{7.1}$$

  En la Figura 7.1, representamos las curvas de autovalores que corresponden al problema semi-discreto (7.1) cuando $f(x) = \sqrt{(2)} * \sin(\pi x/4)$ (izquierda) y $f(x) = \tan(\pi x/4)$ (derecha). Observamos que para ambos casos de mallas no-uniformes, el gap espectral es uniforme. Por consiguiente, esperamos que las propiedades de propagación de las ondas discretas tengan lugar de manera uniforme en el paso del mallado.

- **Diseñar métodos GD que tienen propiedades de propagación apropiadas.** Los métodos SIPG y LDG que estudiamos en esta memoría son unos de los métodos GD diseñados

para problemas elípticos y después utilizados de manera eficiente para aproximar la solución de la ecuación de ondas (cf. [27]). Sin embargo, existen otros métodos GD utilizados en el contexto elíptico. La clase más general de métodos DG son los GD hibridizables introducidos en [21]. Un problema interesante es hacer un análisis de Fourier riguroso de esta clase extendida de métodos GD e identificar los mejores desde el punto de vista de la propagación.

- **Diseñar métodos de típo mixto para las aproximaciones de orden alto de la ecuación de ondas.** Este asunto está relacionado con el tema discutido en el tercer capítulo. Consiste en analizar la eficiencia de los elementos finitos mixtos de orden superior: elementos finitos P^{k_1} para la velocidad y P^{k_2} para la posición, con $k_1 < k_2$. Esta estrategia funciona bien en el caso $(k_1, k_2) = (0, 1)$ para la semidiscretización de la ecuación de ondas $1-d$ y $2-d$ en mallas tanto uniformes como no-uniformes ([17], [18]). Sería interesante hacer un análisis sistemático para valores generales de los parámetros k_1 y k_2.

- **Adaptar la construcción tipo Gaussian beams de paquetes de ondas para sistemas de ecuaciones hiperbólicas continuos y aproximaciones GD.** En el primer capítulo, hemos adaptado la construcción de los Gaussian beams (cf. [46]), muy conocida en el contexto del análisis de propagación de singularidades para problemas hiperbólicos continuos, al caso particular de la semi-discretización por diferencias finitas de la ecuación de ondas d-dimensional. Sería interesante adaptar esta construcción para sistemas de ecuaciones hiperbólicas continuas y para los métodos numéricos más sofisticados como los GD o conformes de orden superior estudiados a lo largo de la tesis.

- **Propiedad de observabilidad para las semi-discretizaciones por métodos GD y elementos finitos clásicos cuadráticos obtenida a través de multiplicadores discretos.** El método habitual para obtener desigualdades de observabilidad para la ecuación de ondas continua consiste en usar el multiplicador $x \cdot \nabla u$ (cf. [36]). Algunas versiones de este multiplicador se han diseñado y usado en el marco discreto (cf. [58]) para las aproximaciones más clásicas de la ecuación de ondas mediante diferencias finitas y elementos finitos clásicos lineales de la ecuación de ondas. Sería interesante construir versiones discretas de este multiplicador y probar su eficacia en el caso de los métodos más sofisticados tratados en esta tesis, tanto GD como elementos finitos clásicos cuadráticos.

- **Propiedades de propagación para las versiones multi-dimensionales de las aproximaciones GD de la ecuación de ondas.** Los métodos GD son mucho más interesantes y complejos en varias dimensiones espaciales, donde la geometría de la malla es mucho más rica. En vista del análisis que hemos desarrollado en el caso $1-d$, sería interesante hacer un análisis de Fourier riguroso de las aproximaciones GD de la ecuación de ondas d-dimensional e identificar los rangos de números de ondas críticos donde los símbolos de Fourier son singulares. Sin embargo, este problema parece bastante complicado desde un punto de vista técnico, porque el número de grados de libertad crece demasiado. Por ejemplo, para el caso más simple de las aproximaciones P_1 en el caso $2-d$ en una malla triangular uniforme, cada punto nodal pertenece a seis triangulos, por tanto hay que usar seis tipos de funciones de base que van a generar seis relaciones de dispersión.

- **Propiedades de propagación de las aproximaciones GD completamente discretas de la ecuación de ondas.** Otro problema abierto consiste en hacer un análisis de Fourier riguroso de las aproximaciones GD completamente discretas de la ecuación de ondas $1-d$ en una malla espacio-temporal uniforme y deducir de allí las consecuencias apropiadas sobre la propagación de ondas.

- **Observabilidad/controlabilidad desde el borde para las aproximaciones por métodos DG y elementos finitos clásicos cuadráticos de la ecuación de ondas $1-d$ en un intervalo acotado**. Consideramos el problema de controlabilidad exacta asociado a las semi-discretizaciones de la ecuación de ondas $1-d$ por elementos finitos clásicos cuadraticos

y métodos DG en un intervalo acotado con un control actuando en uno de los extremos del intervalo espacial y el problema de observabilidad correspondiente al problema adjunto. Un problema interesante consiste en identificar el marco funcional apropiado para aplicar el método HUM (cf. [36]) y ver si basta un único control numérico para controlar los dos modos Fourier (en una clase de datos iniciales filtrada).

- **Construir los controles numéricos que corresponden a los mecanismos de filtrado descritos para elementos finitos clásicos cuadráticos y métodos GD.** Como sabemos, los problemas de controlabilidad exacta y observabilidad del problema adjunto son equivalentes a través del HUM (cf. [36]), de modo que cuando un resultado de observabilidad ocurre de manera uniforme con respecto al paso del mallado dentro de un cierto subespacio de datos iniciales, se tiene también un resultado de controlabilidad exacta de las proyecciones sobre el dual de dicho subespacio, con controles uniformemente acotados con respecto al paso del mallado en norma L^2. Sería interesante encontrar las proyecciones apropiadas que se controlan de manera exacta en el caso de los mecanismos de filtrado que hemos diseñado para las aproximaciones por métodos GD y elementos finitos clásicos cuadráticos de la ecuación de ondas $1-d$.

Chapter 8

Further comments and open problems

In this thesis, we have analyzed the following problems:

1. Constructions of high frequency numerical solutions and their consequences on the divergence rate of the observability constant for the finite-difference approximations of the wave equation.

2. Fourier analysis and filtering mechanisms for two classes of discontinuous Galerkin semi-discretizations of the $1-d$ wave equation (SIPG and LDG).

3. Fourier analysis of the spline semi-discretizations of the $1-d$ wave equation.

4. Fourier analysis and filtering mechanisms for the quadratic classical finite-element semi-discretization of the $1-d$ wave equation.

For the finite-difference approximation of the wave equation, we constructed wave packets for which the observability constant blows-up in a polynomial manner with respect to the mesh size.

For the C^k semi-discretizations, we show that the corresponding group velocity vanishes for $\xi = \pm\pi/h$, allowing to construct solutions propagating arbitrarily slow. Using a Fourier truncation as filtering algorithm, we show that as k increases, the observability time improves with respect to the one corresponding to the P_1-classical finite element.

For the DG or the quadratic classical finite element semi-discretizations, we show that, in view of the Fourier analysis, there exist wave packets propagating arbitrarily slow located on both physical and spurious dispersion diagrams and conclude that the corresponding observability constant blows-up at least polynomially at arbitrary order. We also design some classes of initial data obtained by Fourier truncations or bi-grid algorithms, for which the observability property is recovered uniformly with respect to the mesh size. These filtering mechanisms are twofold: to project the numerical solution on the physical dispersion diagram and to filter out the high frequency numerical solutions concentrated on that diagram.

In what follows, we present in a systematic way some of the open problems that arise in connection with the topics we have addressed and analyzed within this thesis.

- **Exponential divergence for the observability constant.** As we have proved in Proposition 3.2.1 of Chapter 1, by means of a result in [42] concerning the semi-discrete observability problem on an interval of finite length, one can deduce the existence of initial data in (2.1), for which the observability constant $C_h(T)$ in (2.3) blows-up exponentially. But that result does not provide the shape of the corresponding solution. As we have seen, our Gaussian constructions provide a polynomial blow-up order of $C_h(T)$. It would be interesting to further develop our constructions of high-frequency wave packets to prove the exponential blow-up of the observability constant in more general situations.

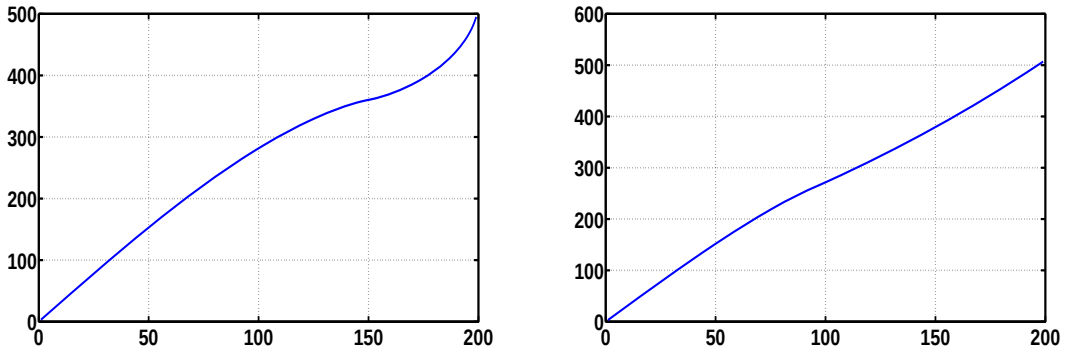

Figure 8.1: Discrete eigenvalues on two particular non-uniform meshes.

- **Propagation properties of the numerical methods on discrete heterogeneous media.** In general, the analysis of propagation and dispersion properties for the approximations of the wave equation is based on Fourier representation of the numerical solution. For numerical schemes with constant coefficients on uniform meshes, this topic is quite well understood. However less is known on this subject on discrete heterogeneous media represented by irregular meshes or variable coefficients, even for the most classical schemes like the finite differences or the classical finite elements method. These are very difficult issues in which even elementary problems like the definition of the symbol of the operator are still open, except for some particular cases obtained by regular deformations from uniform meshes or on periodic grids [10], [23]. Therefore, the adaption of our constructions of highly oscillatory wave packets to this case is a very challenging open problem.

  We also think that it is possible to construct adapted grids, in which the propagation properties of the discrete waves are maintained uniformly with respect to the mesh size. Consider an uniform grid $x_j = jh$ of $(0,1)$, $0 \leq j \leq N+1$ and a non-uniform one y_j obtained as $y_j = f(x_j)$, with $f : (0,1) \to (0,1)$. On the non-uniform grid, we consider the semi-discrete problem:

$$\begin{cases} \partial_{tt}\phi_j(t) - \left(\frac{2(\phi_{j+1}(t)-\phi_j(t))}{(y_{j+1}-y_j)(y_{j+1}-y_{j-1})} - \frac{2(\phi_j(t)-\phi_{j-1}(t))}{(y_j-y_{j-1})(y_{j+1}-y_{j-1})} \right) = 0, & 1 \leq j \leq N \\ \phi_0 = \phi_{N+1} = 0, & \\ \phi_j(0) = \phi_j^0, \quad \partial_t\phi_j(0) = \phi_j^1, & 1 \leq j \leq N. \end{cases} \tag{8.1}$$

  In Figure 8.1, we plot the eigenvalue curves corresponding to the discrete problem (8.1) when $f(x) = \sqrt{(2)} * \sin(\pi x/4)$ (left) and $f(x) = \tan(\pi x/4)$ (right). We observe that on both particular cases of non-uniform meshes, the spectral gap is uniform, therefore we expect observability properties to hold uniformly with respect to the mesh size.

- **Designing DG methods having appropriate propagation properties.** The SIPG and LDG methods are one of the simplest DG methods designed for elliptic problems and after that successfully used for the approximation of the wave equation (cf. [27]). But there exist other DG methods used in the elliptic context. The most general class of DG methods is the so-called hybridizable DG methods introduced in [21]. An interesting open problem is to do a rigorous Fourier analysis of this extended class of DG methods and to identify the better ones from a controllability point of view.

- **Designing mixed-type methods for higher-order approximations.** This issue is related to the topics discussed within Chapter 3. It consists in analyzing the efficiency of mixed-type higher order FEM: P^{k_1}-FEM for the velocity and P^{k_2}-FEM for the position,

with $k_1 < k_2$. This strategy works well in the case $(k_1, k_2) = (0, 1)$ for both $1-d$ and $2-d$ semi-discretizations on uniform grids ([17], [18]). It would be interesting to make a systematic analysis for general values of the parameters k_1 and k_2.

- **Adapt the Gaussian beams construction of wave packets for systems of continuous hyperbolic equations and for DG approximations.** In Chapter 1, we proved that the construction of Gaussian beams [46], well known in the context of the propagation of singularities for the continuous hyperbolic problems, can be also adapted for the particular case of the finite difference semi-discretization of the $d-$dimensional wave equation. It would be interesting to adapt this construction for systems of continuous hyperbolic equations and also for more sophisticated numerical methods like the DG or higher order conformal ones.

- **Observability property for the DG and quadratic classical finite element semi-discretizations of the wave equation by using multiplier techniques.** The usual way to obtain observability inequalities for the continuous wave equation is to use the multiplier $x \cdot \nabla u$ (cf. [36]). Some discrete versions of this multiplier have been also used in the discrete context for the finite-difference and the P_1-classical finite element approximations of the wave equation (cf. [58]). It would be interesting to construct a discrete version of this multiplier for these classes of more sophisticated approximations of the wave equation.

- **Propagation properties for the multi-dimensional versions of DG methods.** The DG methods are even more interesting and complex in higher space dimensions, where the geometry of the mesh is much richer. In view of the analysis done for the $1-d$ case, it would be interesting to make a rigorous Fourier analysis of the *DG*-approximations for the d-dimensional wave equation and to identify the wave numbers where the symbols are singular. Nevertheless, this problem is quite complicated from a technical point of view, due to the fact that the number of degrees of freedom increases very much. Thus, for the simplest case of a P_1-approximation in the $2-d$ case, for each nodal point belonging to six triangles, one has to use six types of basis functions generating six dispersions relations.

- **Propagation properties for fully DG discretizations of the wave equations.** Another open problem is to make a rigorous Fourier analysis of the fully approximation of the $1-d$ wave equation on an uniform spatial and temporal grid and to deduce the appropriate propagation properties.

- **Boundary observability/controllability for the DG and higher order classical FEM approximations of the wave equation.** Consider the exact controllability problem associated to the DG or the quadratic classical finite element semi-discretization of the $1-d$ wave equation on a bounded interval with zero boundary condition imposed at one endpoint and a boundary control acting at the other endpoint and the associated observability problem for the adjoint problem. An interesting problem is to identify the right functional setting in order to apply the HUM method (cf. [36]) and decide if an unique numerical control suffices to control both Fourier modes (within a class of filtered initial data).

- **Construct the controls corresponding to the filtering mechanisms described for the quadratic classical finite element and DG methods.** As we know, by means of the HUM, the controllability problems and the observability of the associated adjoint problems are equivalent (cf. [36]), so that, when an observability result is proved to hold uniformly with respect to h for initial data belonging to some subspace of the Hilbert space for which the adjoint problem is well-posed, an exact controllability result also holds for a suitable projection of the solution of the direct problem on the dual of that subspace, with controls uniformly bounded in L^2. It would be interesting to find the corresponding projections which are exactly controllable for each class of initial data involved in the filtering mechanisms we have designed for the DG and quadratic classical finite element semi-discretizations of the $1-d$ wave equation.

Bibliography

[1] M Ainsworth. Dispersive Behaviour of High Order Discontinuous Galerkin Finite Element Method. *Journal of Computational Physics*, 198(1), 2004.

[2] M. Ainsworth, P. Monk, and W. Muniz. Dispersive and dissipative properties of discontinuous Galerkin finite element methods for the second-order wave equation. *J. Sci. Comput.*, 27(1–3), 2006.

[3] F. Ali Mehmeti. *Nonlinear waves on networks.* Akademie Verlag (Berlin and New York), 1994.

[4] P.F. Antonietti, A. Buffa, and I. Perugia. Discontinuous Galerkin approximation of the Laplace eigenproblem. *Comput. Methods Appl. Mech. Engrg.*, 195(25–28), 2006.

[5] D.N. Arnold. An Interior Penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, 19:742–760, 1982.

[6] D.N. Arnold, F. Brezzi, B. Cockburn, and L.D. Marini. Unified analysis of Discontinuous Galerkin Methods for Elliptic Problems. *SIAM J. Numer. Anal.*, 39:1749–1779, 2002.

[7] C. Bardos, G. Lebeau, and J. Rauch. Sharp sufficient conditions for the Observation, Control and Stabilization of waves form the boundary. *SIAM J. Control and Optimization*, 30:1024–1065, 1992.

[8] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations. *J. Comput. Phys.*, 131:267–279, 1997.

[9] F. Bassi, S. Rebay, G. Mariotti, S. Pdinotti, and M. Savini. A higher-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flow. *2nd European Conference on Turbomachinery Fluid Dynamics and Thermodinamics (Antwerpen, Belgium), Technologisch Instituut, March 5-7*, pages 99–108, 1997.

[10] B. Beckermann and S. Serra-Capizzano. On the asymptotic spectrum of finite element matrix sequences. *SIAM J. Numer. Anal.*, 45(2):746–769, 2007.

[11] T. Belytschko and R. Mullen. On dispersive properties of finite element solutions. *Modern Problems in Elastic Wave Propagation, Wiley*, 1978.

[12] J.B. Bowles and R. Vichnevetsky. *Fourier Analysis of Numerical Approximations of Hyperbolic Equations.* SIAM Philadelphia, 1982.

[13] F. Brezzi, B. Cockburn, L.D. Marini, and E. Süli. Stabilization mechanisms in discontinuous Galerkin finite element methods. *Comput. Methods Appl. Mech. Engrg.*, 195:3293–3310, 2006.

[14] F. Brezzi, M. Manzini, L.D. Marini, P. Pietra, and A. Russo. Discontinuous Galerkin approximations for elliptic problems. *Numer. Methods for Partial Differential Equations*, 16:365–378, 2000.

[15] L. Brillouin. *Wave Propagation and Group Velocity.* Academic Press, 1960.

[16] P. Castillo, B. Cockburn, I. Perugia, and D. Schotzau. An a Priori Error Analysis of the Local Discontinuous Galerkin Method for Elliptic Problems. *SIAM J. Numer. Anal.*, 38(5):1676–1706, 2001.

[17] C. Castro and S. Micu. Boundary controllability of a linear semi-discrete $1-d$ wave equation derived from a mixed finite element method. *Mumerische Mathematik*, 102(3):413–462, 2006.

[18] C. Castro, S. Micu, and A. Münch. Numerical approximation of the boundary control for the wave equation with mixed finite elements in the square. *IMA J. of Numerical Analysis*, 28(1):186–214, 2008.

[19] M. Chenais and Zuazua E. Controllability of an elliptic equation and its finite difference approximation by the shape of the domain. *Numerische Mathematik*, 95:63–99, 2003.

[20] C.K. Chui. *An introduction to wavelets.* Academic Press, 1992.

[21] B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.*, 47(2):1319–1365, 2009.

[22] B. Cockburn and C.W. Shu. The local discontinuous Galerkin finite element method for convection-diffusion systems. *SIAM J. Numer. Anal.*, 35:2440–2463, 1998.

[23] S. Ervedoza. Observability properties of a semi-discrete 1d wave equation derived from a mixed finite element method on nonuniform meshes. *ESAIM:COCV, in press.*

[24] S. Ervedoza and E. Zuazua. *Propagation, observation and numerical approximation of waves.* in preparation.

[25] L.C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics.* American Mathematical Society, New York, 1998.

[26] R. Glowinski. Ensuring well posedness by analogy: Stokes problem and boundary control for the wave equation. *J. Comput. Physics*, 103(2):189–221, 1992.

[27] M.J. Grote, A. Schneebeli, and D. Schötzau. Discontinuous Galerkin finite element method for the wave equation. *SIAM J. Numer. Anal.*, 44(6), 2006.

[28] T.R. Hill and W.H. Reed. Triangular mesh methods for the neutron transport equation. *Tech. Report LA-UR-73-479, Los Alamos Scientific Laboratory*, 1973.

[29] T.R.J. Hughes, A. Reali, and G. Sangalli. Duality and unified analysis of discrete approximations in structural dynamics and wave propagation: Comparison of p-method finite elements with k-method NURBS. *Computer Methods in Applied Mechanics and Engineering*, 197(49-50):4104–4124, 2008.

[30] L. Ignat. *Propiedades cualitativas de esquemas numéricos de aproximación de ecuaciones de difusión y de dispersión.* PhD thesis, Universidad Autónoma de Madrid, 2006.

[31] L. Ignat and E. Zuazua. Dispersive properties of numerical schemes for nonlinear Schrödinger equations. *Foundations of computational mathematics FoCM Proceedings, Santander 2005*, 331:181–207, 2006.

[32] L. Ignat and E. Zuazua. Convergence of a two-grid algorithm for the control of the wave equation. *J.Eur.Math.Soc.*, 11(2):351–391, 2009.

[33] J.A. Infante and E. Zuazua. Boundary observability for the space-discretizations of the one-dimensional wave equation. *M2AN*, 33(2):407–438, 1999.

[34] G. Lebeau. Contrôle de l'équation de Schrödinger. *J. Math. Pures Appl.*, 71:267–291, 1992.

[35] G. Lebeau and E. Zuazua. Decay rates for the three-dimensional linear system of thermoelasticity. *Archives Rat. Mech. Anal.*, 148:179–231, 199.

[36] J.-L. Lions. *Contrôlabilité exacte, perturbations et stabilization de systèmes distribués. Tome 1: Contrôlabilité exacte*. Masson, 1988.

[37] F. Macià. *Propagación y control de vibraciones en medios discretos y continuos*. PhD thesis, Universidad Complutense de Madrid, 2002.

[38] F. Macià and E. Zuazua. On the lack of observability for wave equations: a Gaussian beam approach. *Asymptotic Analysis*, 32:1–26, 2002.

[39] A. Marica and E. Zuazua. Localized waves for the finite-difference semi-discretization of the wave equation. *V International Conference on Inverse Problems, Control and Shape Optimization, PICOF'10 Proceeding, April 7-9, 2010, Cartagena, Spain.*

[40] A. Marica and E. Zuazua. Localized solutions for the finite difference semi-discretization of the wave equation. *C.R.Acad.Sci. Paris*, 348(11–12):647–652, 2010.

[41] A. Marica and E. Zuazua. Localized solutions and filtering mechanisms for the discontinuous Galerkin semi-discretizations of the 1-d wave equation. *C.R.Acad.Sci. Paris*, submitted.

[42] S. Micu. Uniform boundary controllability of a semi-discrete 1-D wave equation. *Numer. Math*, 91(4):723–768, 2002.

[43] S. Micu. Uniform boundary controllability of a semi-discrete 1-d wave equation with vanishing viscosity. *SIAM J. Control Optim.*, 47(6):2857–2885, 2008.

[44] M. Negreanu. *Métodos numéricos para el anàlisis de la propagación, observación y control de ondas*. PhD thesis, Universidad Complutense de Madrid, 2003.

[45] M. Negreanu and E. Zuazua. Convergence of a multigrid method for the controllability of a $1-d$ wave equation. *C.R.Math.Acad.Sci.Paris*, 338:413–418, 2004.

[46] J. Ralston. Gaussian beams and the propagation of singularities. *in Studies in Partial Differential Equations, ed. by W. Littman, MAA Studies in Mathematics, Mathematical Association of America*, 23:206–248, 1983.

[47] J. Ralston. *Gaussian Beams*. http://www.math.ucla.edu/ ralston/pub/Gaussnotes.pdf, September 2005.

[48] J. Rauch, X. Zhang, and E. Zuazua. Polynomial decay for a hyperbolic-parabolic coupled system. *J. Math. Pures Appl.*, 84:407–470, 2005.

[49] I.J. Schoenberg. *Cardinal spline interpolation*, volume 12 of *CBMS Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, 1973.

[50] E. Stein. *Harmonic Analysis: Real variable methods, orthogonality and oscillatory integrals*. Princeton University Press, New Jersey, 1993.

[51] J.C. Strikwerda. *Finite difference schemes and partial differential equations*. SIAM, 2004.

[52] L.R. Tcheugoué Tébou and E. Zuazua. Uniform exponential long time decay for the space semi-discretizations of a damped wave equation with artificial numerical viscosity. *Numerische Mathematik*, 95(3):563–598, 2003.

[53] L. Trefethen. *Finite Difference and Spectral Methods for Ordinary and Partial Differential Equations*. http://www.comlab.ox.ac.uk/nick.trefethen/pdetext.html, 1994.

[54] L.N. Trefethen. Group Velocity in Finite Difference Schemes. *SIAM Review*, 24(2):113–136, 1982.

[55] G.B. Whitham. *Linear and Nonlinear Waves*. Wiley-Interscience, John Wiley & Sons, New York-London-Sydney, 1974.

[56] X. Zhang and E. Zuazua. Polynomial decay and control of a $1-d$ hyperbolic-parabolic coupled system. *J. Differential Equations*, 15:205–235, 1990.

[57] E. Zuazua. Exponential decay for the semilinear wave equation with localized damping in unbounded domains. *J. Math. Pures Appl.*, 70:513–529, 1991.

[58] E. Zuazua. Boundary Observability for the Finite-Difference space semi-discretizations of the $2-d$ wave equation in the square. *J. Math. Pures Appl.*, 78:523–563, 2002.

[59] E. Zuazua. Propagation, Observation, Control and Numerical Approximations of Waves. *SIAM Review*, 47(2):197–243, 2005.

[60] E. Zuazua. *Handbook of Differential Equations: Evolutionary Equations*, volume 3, chapter Controllability and Observability of Partial Differential Equations: Some results and open problems, pages 527–621. Elsevier Science, 2006.