Available online at www.sciencedirect.com

ScienceDirect

Journal of Differential Equations 309 (2022) 949–993

Journal of Differential Equations

www.elsevier.com/locate/jde

A multi-scale Gaussian beam parametrix for the wave equation: The Dirichlet boundary value problem

Michele Berra [a,1], Maarten V. de Hoop [b], José Luis Romero [c,d,*]

[a] *Buzzi Unicem S.p.A., Italy*

[b] *Department of Computational and Applied Mathematics, Rice University, 6100 Main MS-134, Houston, TX 77005, United States of America*

[c] *Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria*

[d] *Acoustics Research Institute, Austrian Academy of Science, Wohllebengasse 12-14, 1040 Vienna, Austria*

Received 5 December 2018; accepted 24 November 2021

Available online 8 December 2021

Abstract

We present a construction of a multi-scale Gaussian beam parametrix for the Dirichlet boundary value problem associated with the wave equation, and study its convergence rate to the true solution in the highly oscillatory regime. The construction elaborates on the wave-atom parametrix of Bao, Qian, Ying, and Zhang and extends to a multi-scale setting the technique of Gaussian beam propagation from a boundary of Katchalov, Kurylev and Lassas.

MSC: 35L05; 35L20; 35S05; 42C15

Keywords: Gaussian beam; Wave equation; Wave-atom; Parametrix; Boundary-value problem

* Corresponding author at: Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria.

E-mail addresses: michele.berra@polito.it (M. Berra), mdehoop@rice.edu (M.V. de Hoop), jose.luis.romero@univie.ac.at, jlromero@kfs.oeaw.ac.at (J.L. Romero).

[1] Formerly at Politecnico di Torino (DISMA), corso Duca degli Abruzzi 24, 10129 Torino, Italy.

https://doi.org/10.1016/j.jde.2021.11.043

0022-0396/

1. Introduction

1.1. The parametrix

Gaussian beams (GB) are high-frequency asymptotic solutions for hyperbolic partial differential equations, in particular, for the homogeneous wave equation,

$$u_{tt}(t,x) - c(x)^2 \Delta_x u(t,x) = 0. \tag{1.1}$$

Gaussian beams are initiated via an Ansatz. They follow the propagation of singularities, that is, the bicharacteristics associated with the principal symbol of the wave operator, while distinguishing themselves from standard geometrical optics solutions by capturing the asymptotic behavior in caustics without precautions; see Fig. 2.3.

GB parametrices for the initial value problem (IVP) for the wave equation are based on representations of the initial data as superimposition of certain Gaussian-like wavepackets:

$$u(0,x) = \sum_\gamma a_\gamma \varphi_\gamma(x), \qquad u_t(0,x) = \sum_\gamma b_\gamma \varphi_\gamma(x). \tag{1.2}$$

Each wavepacket φ_γ is then used to generate two GB: $\varphi_\gamma(x) \approx \Phi_\gamma^\pm(0,x)$, where the choice of sign $\pm$ corresponds to the two polarized modes of the wave-equation. Specifically, one constructs a frame of such wavepackets that initialize multi-scale Gaussian beams, and the resulting parametrix has the form

$$\tilde{u}(t,x) = \sum_\gamma \alpha_\gamma^+ \Phi_\gamma^+(t,x) + \sum_\gamma \alpha_\gamma^- \Phi_\gamma^-(t,x), \tag{1.3}$$

where the coefficients $\alpha_\gamma^\pm$ are defined in terms of a_γ and b_γ. The precise form of the packet decomposition in (1.2) determines the effectiveness of the parametrix. A detailed study of the approximation error of such parametrices when the initial data is a finite sum of Gaussian packets is provided in [28]. Gaussian-beam parametrices and summation of Gaussian beams are naturally connected to Fourier integral operators with complex phase.

Several other parametrices for the wave equation are also based on wavepacket expansions. Indeed, Smith [43,44] introduced the use of a frame of wavepackets with parabolic scaling (curvelets) in the construction of a parametrix, which, for smooth wave speeds, can be identified as a Fourier integral operator. This representation is also underlying the analysis of wave propagators of Candès and Demanet [7]. Further related constructions based on localized wavepackets can be found in the work of Tataru [48], Koch and Tataru [24], Geba and Tataru [18], and De Hoop, Uhlmann, Vasy and Wendt [14].

In [40,4] a GB parametrix was introduced where the beams are initialized following the wave-atom tiling of phase space [15,16]. Thus, the frequency profile of the initial Gaussian packets is adapted to the cover depicted in Fig. 2.1. (See also [49].) The merit of using wave atoms is that they are both *isotropic* - as required in order to apply the GB method - and *parabolic* - in the sense that their frequency center ξ and the diameter of their essential frequency support ℓ satisfy $\ell^2 \approx |\xi|$. The resulting parametrix has order $1/2$, performing similarly to the ones based on curvelets with second-order corrections [44,43,7,12].

In this paper, we introduce a GB parametrix for the Dirichlet boundary value problem (BVP) associated with the wave equation and analyze its approximation properties. For simplicity, we assume that the boundary is flat and treat the model case of the half space $\mathbb{R}^d_+ = \{x \in \mathbb{R}^d : x_1 > 0\}$,

$$\begin{cases} u_{tt}(t,x) - c(x)^2 \Delta_x u(t,x) = 0, & t \in [0,T], x \in \mathbb{R}^d_+, \\ u(0,x) = u_t(0,x) = 0, & x \in \mathbb{R}^d_+, \\ u(t,0,y) = h(t,y), & t \in [0,T], y \in \mathbb{R}^{d-1}, \end{cases} \tag{1.4}$$

with c being smooth and bounded below by a positive constant, and h being prescribed. In the applications to reverse-time continuation from the boundary, as it appears in imaging, for example, h represents boundary data on an acquisition manifold $\{0\} \times \mathbb{R}^{d-1}$ with time interval $[0, T]$.

We consider a wave-atom like expansion of the boundary value,

$$h(t,y) = \sum_\gamma h_\gamma \varphi_\gamma(t,y), \tag{1.5}$$

and construct adequate Gaussian beams $\Phi^\pm_\gamma$, so that they match the wavepackets along the boundary:

$$\Phi^\pm_\gamma(t,0,y) \approx \varphi_\gamma(t,y). \tag{1.6}$$

As parametrix solution for the Dirichlet problem we then propose:

$$\tilde{u}(t,x) = \sum_\gamma h_\gamma \Phi^\pm_\gamma(t,x). \tag{1.7}$$

Based on the effectiveness of the parametrix for the IVP, the expectation is that u be an approximate solution for the homogeneous wave equation. The beams Φ_γ have to be designed with the additional requirement that at initial time the parametrix and its time derivative be approximately null: $\tilde{u}(0,x)$, $\tilde{u}_t(0,x) \approx 0$. With this provision, the energy estimates [27,26] imply that the parametrix solution is close to the true one.

A key application of the Gaussian beam method is imaging in reflection seismology [6,19] - see also [37] for an analysis of imaging and its connection with solving boundary value problems. There is extensive work done on computations with Gaussian beams and wavepackets [47,34,2,41]. We expect these to be instrumental to the implementation of the parametrix that we introduce, thus facilitating accurate computations in the presence of caustics.

We now elaborate on the details of the program for the construction and analysis of the parametrix outlined above.

(i) Description of boundary restriction of beams. At an initial time, Gaussian beams have a prescribed Gaussian profile on x. The GB theory provides estimates for the evolution of this profile for subsequent times t. In contrast, the approximation in (1.6) requires describing the restriction of a GB to the boundary $\{x_1 = 0\}$ *treating the remaining variables* $(t, x_2, \dots, x_d)$ *jointly as a spatial variable*. Such an analysis is the first step of our construction: We consider a general Gaussian beam and approximately describe its restriction to the acquisition manifold as a Gaussian wavepacket in all remaining variables including time. This elaborates on a technique

of Katchalov, Kurylev and Lassas [22], who considered the boundary restriction of a Gaussian beam that intersects the boundary through a normal ray.

(ii) Packet-beam matching. Given a general (isotropic) Gaussian wavepacket, φ_γ, we use the analysis from (i) to construct an adequate beam satisfying (1.6). This defines a map $\mathcal{S}$ that assigns to every phase-space parameter γ indexing the packets in the expansion of the boundary value h (1.5) a set of initial conditions $\mathcal{S}_\gamma$ for the ordinary differential equations (ODE) that define a Gaussian beam.

(iii) Back-propagation. The packet-beam matching (ii) is carried out as follows: given a wavepacket $\varphi_\gamma(t, y)$ with spatial center (t_γ, y_γ), we construct the beam $\Phi_\gamma(t, x)$ so that its spatial center intersects the boundary precisely at time $t = t_\gamma$. The profile of the beam is specified at time $t = t_\gamma$ and back-propagated to time $t = 0$ by means of the defining ODEs. Additionally, we specify the mode of Φ_γ - that determines in which direction bicharacteristics are traveled - so that the beam moves *into* the half-space as time evolves. As a consequence the beam Φ_γ is mostly concentrated outside the right half-space at $t = 0$, and the parametrix approximately vanishes at initial time, as required in order to apply energy estimates.

(iv) Distortion of the phase-space tiling. Wavepacket expansions such as (1.2) and (1.5) follow certain tilings in phase space. For wave-atom expansions, the frequency variable is partitioned as shown in Fig. 2.1 and the space variable is resolved following the dual scaling. The IVP parametrix relies on this pattern: The technical results in [4] show that for subsequent times the beams $\Phi^\pm(t, \cdot)$ in (1.3) are still adapted to a similar phase-space tiling, and thus enjoy similar spanning properties. While the expansion of the boundary value in (1.5) fits the framework of wave atoms, the phase-space tiling governing the profile of the beams in the proposed parametrix (1.7) is impacted by the packet-beam matching procedure (iii). The analysis of the approximation error of the parametrix involves a careful quantification of this effect.

1.2. Assumptions and results

We assume that the wave speed c is smooth, bounded below by a positive constant and has globally bounded derivatives of every order. (The smoothness assumptions could be relaxed at the cost of a more technical presentation.) The essential condition for the effectiveness of the parametrix that we introduce is that the rays of the associated Hamiltonian that take off from the boundary do not return to the boundary in the time interval in question, so that the back-propagation step (iii) succeeds (see Section 4.1.1 for a precise quantitative formulation).

Besides the standard compatibility condition $h(0, \cdot) = 0$, we also assume that the wavefront set of the boundary value h does not contain grazing rays. While this assumption is not necessary for the Dirichlet problem to be well-posed, our parametrix is ultimately based on oscillatory integrals and the theory of elliptic boundary value problems, and these techniques do require that the bicharacteristic be nowhere tangent to the boundary [36]. We enforce these assumptions by examining the wavepacket expansion of h (1.5) and by discarding (or down-weighting) those coefficients that correspond to the undesired wavefront components. In order to describe this operation in intrinsic terms (i.e., independently of the particular wavepacket expansion that the parametrix uses) we introduce a pseudodifferential cut-off σ that eliminates grazing rays and consider a modified Dirichlet problem with boundary condition $u(t, 0, y) = h_{cut}(t, y) := \sigma(t, y, D_t, D_y)h(t, y)$. Denoting by u the solution of the modified problem, we show that our parametrix solution $\tilde{u}$ satisfies:

$$\|\tilde{u} - u\|_{C^0([0,T],H^1(\mathbb{R}^d_+))\cap C^1([0,T],L^2(\mathbb{R}^d_+))} \leq C_T \|h\|_{H^{1/2}(\mathbb{R}^d)}.$$

In particular, in the highly oscillatory regime, $\hat{h}(\xi) = 0$ for $|\xi| \leq \xi_{\min}$, the error can be estimated in terms of the scale content of the initial data, giving a bound $\xi_{\min}^{-1/2} \cdot \|h\|_{H^1(\mathbb{R}^d)}$.

We also note that the Dirichlet problem is related to a boundary source problem. Let u^r be the solution to (1.4) and u^l the solution to the analogous problem on the left-half space $\mathbb{R}^d_- = \{x \in \mathbb{R}^d : x_1 < 0\}$. Let

$$u(t,x) := u^r(t,x)H(x_1) + u^l(t,x)H(-x_1), \tag{1.8}$$

where H denotes the Heaviside function. Then (1.8) is a microlocal solution to a boundary-source problem:

$$u_{tt}(t,x) - c(x)^2\Delta_x u(t,x) = (Bh)(t,x_*)\delta_{x_1}(x), \tag{1.9}$$

where B is an adequate boundary operator [37,46]. Hence, our parametrix provides also a microlocal solution to (1.9).

1.3. Related work

The construction of Gaussian beams dates back to the 1960's, that is, the work by Babič and Buldyrev [3].[2] Later, Gaussian beams were used in the analysis of regularity and propagation of singularities in partial differential and pseudodifferential equations by Hörmander [20] and Ralston [42]. Without any attempt to give a comprehensive list of references, we refer to the foundational work of Popov [38,39] and Katchalov and Popov [21], and the applications to seismic wave propagation by Červenỳ, Popov and Pšenčík [8]. Furthermore, we mention connections with complex rays in the work of Keller and Streifer [23], and Deschamps [17] in the early 1970s, and the work of Weston [50], who studied the wave splitting in a flat boundary, which is part of the parametrix construction for boundary value problems.

Our study of the Dirichlet problem builds fundamentally on the work of Katchalov, Kurylev and Lassas [22], who describe the boundary restriction of a single normally incident Gaussian beam. We extend this analysis to a collection of multi-scale Gaussian beams with varying incidence angles.

Wave parametrices based on Gaussian wavepacket expansions go back to Córdoba and Fefferman [10], and related techniques can be found, for example, in the work of Smith [43,44], Candès and Demanet [7], Tataru [48], Koch and Tataru [24], and Geba and Tataru [18]. In the context of Gaussian beams, Liu, Runborg and Tanushev studied convergence rates of parametrices for initial data consisting of a finite sum of Gaussian wavepackets [28]. Our construction elaborates particularly on the work of Bao, Qian, Ying, and Zhang [41,4] who treat the decomposition of general (multi-scale) initial data. Indeed, much of the technical work in this article is devoted to show that the packet-beam matching procedure described above yields a family of beams that approximately resemble at an initial time the wavepackets used in [4] as a starting point for the IVP. This task leads us to introduce the notion of *well-spread family of Gaussian beams*, that abstracts the properties that make a multi-scale GB parametrix effective. We also mention a link of our analysis with the work of Laptev and Sigal [25] who constructed a parametrix for the time-dependent Schrödinger equation.

[2] The book of Babič and Buldyrev was translated; it contains work that Babich and his colleagues published in the Proceedings of the Steklov Institute in 1968.

While the no-grazing ray assumption is standard in the literature, we note that the existence and uniqueness theorems for initial-boundary value problems do not involve these transversality conditions, and, indeed, Melrose [31] introduced a class of operators to treat glancing points, and used them in a parametrix construction [32]; see also [33].

In relation to Gaussian beam expansions, we also mention the related notion of frozen Gaussian beams [30,29], where the Gaussian packets that approximate the solution to the wave equation may themselves not be asymptotic solutions.

1.4. Organization

In Section 2 we introduce multi-scale Gaussian beams and a frame of wave-atom like Gaussian wavepackets. We also discuss how to parametrize GB by their initial conditions, introduce the relevant notation, and collect some facts about the defining ODEs. The notion of well-spread family of Gaussian beams is introduced in Section 3. We show that such families enjoy suitable uniformity properties, satisfy Bessel bounds and are approximate solutions to the homogeneous wave equation. In Section 4 we introduce the Dirichlet BVP and the corresponding assumptions. In Section 5 we analyze a family of beams at times where their spatial centers intersect the acquisition manifold. This is then used as a guide in Section 6 to introduce the beam-packet matching procedure. The most technical proofs are postponed to Section 8. The performance of the parametrix is finally analyzed in Section 7. In some cases, we skip or only sketch the proof of certain technical lemmas. The reader can consult our technical report [5] for full details. More notation is introduced throughout the paper; a reference table can be found in Appendix 9.

1.5. Notation

We write $x = (x_1, x_*) \in \mathbb{R} \times \mathbb{R}^{d-1}$, $|x| = |x|_2$ denotes the Euclidean norm, $\mathbb{R}^d_+ = (0, +\infty) \times \mathbb{R}^{d-1}$, and $\mathbb{R}^d_T = [-T, T] \times \mathbb{R}^{d-1}$. We use the notation $B_r(x)$ for the Euclidean ball of center x and radius r. $\Re(z)$ and $\Im(z)$ denote respectively the real and imaginary part of $z \in \mathbb{C}$. This notation extends to vectors and matrices componentwise. Generic constants are denoted by C, C', C_0 and their meaning may change from line to line. Specific constants are given more descriptive notation.

For two non negative functions f, g, $f \lesssim g$ means that there exists a constant $C > 0$ such that $f(x) \le C g(x)$, for all x. We write $f \asymp g$ if $f \lesssim g$ and $g \lesssim f$. Given a domain $\Omega \subseteq \mathbb{R}^d$, we let $C_b^\infty(\Omega)$ be the class of $C^\infty(\Omega)$ functions f such that for every multi-index α, $\partial_x^\alpha f \in L^\infty(\Omega)$.

The identity matrix is denoted as $I_d \in \mathbb{R}^{d\times d}$. For a matrix $A \in \mathbb{C}^{d\times d}$, $A \gtrsim I_d$ means that there exists a constant $C > 0$ such that $A - C \cdot I_d$ is a positive matrix (i.e. Hermitian and with non-negative spectrum). For a constant $C \ge 0$, we sometimes write $A \ge C$ instead of $A \ge C I_d$. The Fourier transform is normalized as: $\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i x\xi} dx$. The phase-space metric is the function $d : \mathbb{R}^d \times \left(\mathbb{R}^d \setminus \{0\}\right) \to [0, +\infty)$,

$$d((x,\xi),(x',\xi')) = |\xi||\xi'||x - x'|^2 + \left|\xi - \xi'\right|^2 \qquad (x,\xi),(x',\xi') \in \mathbb{R}^{2d}.$$

The Hamiltonians are defined as $H^+(x, p) = c(x)\,|p|$, $H^-(x, p) = -c(x)\,|p|$ and H denotes generically either H^+ or H^-. Sometimes we denote time derivatives with a dot, e.g. $\dot{x}(t) = \partial_t x(t)$.

Throughout the article, c denotes a fixed function $c \in C_b^\infty$ (called velocity) that is assumed to be bounded below away from 0; i.e.,

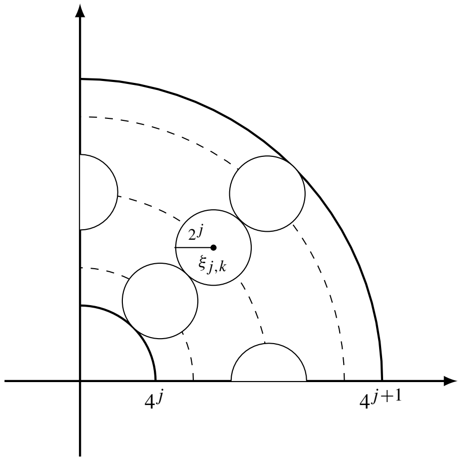

Fig. 2.1. Frequency-space tiling for the frame.

$$C_{\text{vel}} := \inf_{x \in \mathbb{R}^d} c(x) > 0, \tag{1.10}$$

and $\partial_x^\alpha c \in L^\infty(\mathbb{R}^d)$ for every multi-index α.

2. Gaussian wavepackets and Gaussian beams

2.1. The frame

We construct a frame of wave-atom-like Gaussian packets with Gaussians as basic waveforms. We start by introducing a frequency cover. For $j \geq 1$ we let $\left\{\xi_{j,k} : k = 1, \ldots, n_j\right\} \subseteq \mathbb{R}^d$ be a set of points such that:

- The family $\left\{B_{2^j}(\xi_{j,k}), k = 1, \ldots, n_j\right\}$ is disjoint and each member is contained in the corona $\mathcal{C}_j = B_{4^{j+1}}(0) \setminus B_{4^j}(0)$.
- $\mathcal{C}_j \subseteq \bigcup_{k=1}^{n_j} B_{2^{j+1}}(\xi_{j,k})$.

Hence $\{B_{2^{j+1}}(\xi_{j,k}) : k = 1, \ldots, n_j, j \geq 1\}$ is a cover of $\{\xi \in \mathbb{R}^d : |\xi| \geq 4\}$; see Fig. 2.1.

Note that $\left|\xi_{j,k}\right| \asymp 4^j$. In addition, comparing the volumes of $\mathcal{C}_j$ to those of the unions of the balls $B_{2^j}(\xi_{j,k})$ and $B_{2^{j+1}}(\xi_{j,k})$, it follows that $n_j \asymp 2^{jd}$. For convenience, we also introduce the rescaled vector

$$\tilde{\xi}_{j,k} = 2\pi \frac{\xi_{j,k}}{4^j}. \tag{2.1}$$

Hence, $\tilde{\xi}_{j,k}$ is approximately normalized: $|\tilde{\xi}_{j,k}| \asymp 1$.

We let $\varphi(x)$ be the Gaussian function

$$\varphi(x) = 2^{\frac{d}{4}} e^{-\pi|x|^2}, \qquad x \in \mathbb{R}^d, \tag{2.2}$$

and define modulated and scaled waveforms adapted to the frequency cover

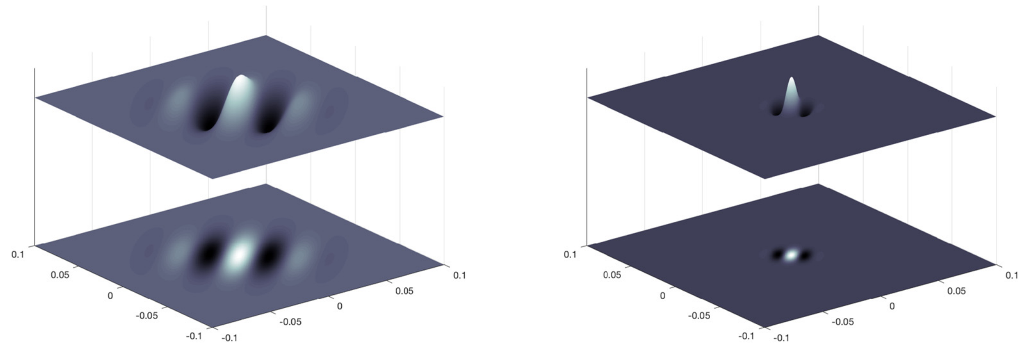

Fig. 2.2. On the left, the space profile and contour plot of a frame element with $j = 2$. On the right, the same plot with a Gaussian window contracted by a factor of $2\pi \ln(16)$, which is better adapted for certain numerical examples.

$$\varphi_{j,k}(x) = 2^{j\frac{d}{2}} e^{2\pi i \xi_{j,k} x} \varphi(2^j x), \quad j \geq 1, k = 0, \dots, n_j,$$

so that $\widehat{\varphi_{j,k}}$ is essentially concentrated on $B_{2^j}(\xi_{j,k})$. We let $\Lambda \subseteq \mathbb{R}^d$ be a (full rank) lattice and define

$$\Gamma = \left\{ (j, k, \lambda) : j \geq 1, k = 0, \dots, n_j, \lambda \in \Lambda \right\}, \tag{2.3}$$

and

$$\varphi_\gamma(x) = \varphi_{j,k,\lambda}(x) = \varphi_{j,k}(x - 2^{-j}\lambda), \qquad \gamma = (j, k, \lambda) \in \Gamma.$$

Explicitly,

$$\varphi_{j,k,\lambda}(x) = 2^{j\frac{d}{2}} e^{2\pi i \xi_{j,k}(x - 2^{-j}\lambda)} \varphi(2^j x - \lambda), \qquad (j, k, \lambda) \in \Gamma.$$

See Fig. 2.2 for a plot. For an index $\gamma \in \Gamma$, we often refer implicitly to the notation $\gamma = (j, k, \lambda)$.

Although we are mainly interested in high frequency expansions: $\sum_{\gamma \in \Gamma} f_\gamma \varphi_\gamma$, in order to expand an arbitrary function we need to provide wavepackets adapted to the zeroth-scale. To keep the notation concise, we let $\varphi_{0,0} := \varphi$, augment the index set Γ by

$$\Gamma_* := \Gamma \cup \{(0, 0, \lambda) : \lambda \in \Lambda\}, \tag{2.4}$$

and define zeroth-scale wavepackets as: $\varphi_{0,0,\lambda} := \varphi(x - \lambda)$. The complete set of wavepackets can be written as:

$$\mathcal{F} = \left\{ \varphi_\gamma : \gamma \in \Gamma_* \right\}.$$

We now show that the system thus constructed is indeed rich enough to represent any function.

Theorem 2.1. *For an adequate lattice $\Lambda \subseteq \mathbb{R}^d$, the system $\mathcal{F}$ is a frame for the inhomogeneous Sobolev spaces $H^s(\mathbb{R}^d)$, with $-1 \leq s \leq 1$. More precisely, the frame operator $S_\mathcal{F} f =$*

$\sum_{\gamma\in\Gamma_*}\langle f,\varphi_\gamma\rangle\varphi_\gamma$ is invertible on $H^s(\mathbb{R}^d)$ for $-1\le s\le 1$. As a consequence, every $f\in H^s(\mathbb{R}^d)$ can be represented by an H^s-convergent series

$$f=\sum_{\gamma\in\Gamma_*} f_\gamma\varphi_\gamma,\qquad f_\gamma:=\left\langle f, S_{\mathcal{F}}{}^{-1}\varphi_\gamma\right\rangle, \tag{2.5}$$

and the following norm equivalences hold

$$\|f\|_{H^s}^2\asymp\sum_{\gamma\in\Gamma_*}4^{2js}\left|\langle f,\varphi_\gamma\rangle\right|^2\asymp\sum_{\gamma\in\Gamma_*}4^{2js}\left|f_\gamma\right|^2.$$

We remark that in Theorem 2.1, the symbol $\langle f,\varphi_\gamma\rangle$ denotes the standard L^2 inner product. The theorem is proved using a variant of Daubechies's criterion for wavelets. Details can be found in our technical report [5] - see also [1,35,11] for related estimates. Let us mention that the construction provides a concrete criterion to choose the lattice Λ and $A\|f\|_{H^s}\le\|S_{\mathcal{F}}f\|_{H^s}\le B\|f\|_{H^s}$ can be satisfied with B/A reasonably small. Hence, the numerical inversion of $S_{\mathcal{F}}$ is well-conditioned.

From now on we fix a lattice Λ such that the conclusion of Theorem 2.1 holds.

2.2. *Operating on the frame expansion*

We will be mostly interested in the higher scales $j\ge 1$. We can truncate the representation in (2.5),

$$\tilde f=\sum_{\gamma\in\Gamma} f_\gamma\varphi_\gamma, \tag{2.6}$$

and it is easy to see that the error can be bounded as $\|f-\tilde f\|_{H^1}\lesssim\|f\|_{H^{-1}}$. Hence, in the highly oscillatory regime, we only need to consider expansions of the form (2.6).

More generally, we use pseudodifferential cut-offs to operate microlocally on a function f and we wish to approximately implement those operations by acting directly on the expansion in (2.6). We recall that a symbol $\sigma:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{C}$ belongs to the Hörmander class $S^0_{1,0}(\mathbb{R}^d\times\mathbb{R}^d)$ if $\left|\partial_x^\beta\partial_\xi^\alpha\sigma(x,\xi)\right|\le C_{\alpha,\beta}(1+|\xi|)^{-|\alpha|}$, for all multi-indices α,β. The next lemma will be an important technical tool. See [5, Appendix C] for a proof, and [13, Lemma 3.1] and [12, Lemma 17] for related results for curvelets.

Theorem 2.2. *Let $\sigma\in S^0_{1,0}(\mathbb{R}^d\times\mathbb{R}^d)$. Then, for $s\in[1/2,1]$ and $f\in H^s(\mathbb{R}^d)$,*

$$\|\sigma(x,D)f-\sum_{\gamma\in\Gamma}\sigma(2^{-j}\lambda,\xi_{j,k})f_\gamma\varphi_\gamma\|_{H^s}\lesssim\|f\|_{H^{s-1/2}}.$$

(Here, $\sigma(x,D)$ is the Kohn-Nirenberg quantization of σ, and $f_\gamma:=\langle f,S_{\mathcal{F}}{}^{-1}\varphi_\gamma\rangle$ are the high-scale frame coefficients of f.)

2.3. Gaussian beams

We summarize the construction of Gaussian beams, following the treatment of Katchalov, Kurylev and Lassas [22]. We seek formal asymptotic solutions to the wave equation 1.1 (in a “moving” frame of reference) of the form

$$\Phi(t,x) = A(t)e^{i\omega\theta(t,x)},$$

where the phase function θ and amplitude function A are smooth and complex-valued, and ω is the frequency parameter. Asymptotic analysis of the eikonal and transport equations leads a second order phase-function

$$\theta(t,x) = p(t)\cdot(x-x(t)) + \tfrac{1}{2}(x-x(t))^T M(t)(x-x(t)),$$

with the following ingredients. First, the analysis of propagation of singularities dictates that $x^{\pm}(t)$, $p^{\pm}(t)$ be described by bicharacteristics, satisfying the Hamilton system

$$\dot{x}^{\pm}(t) = \quad \partial_p H^{\pm}, \quad \dot{p}^{\pm}(t) = -\partial_x H^{\pm}, \tag{2.7}$$

supplemented with initial conditions, $x^{\pm}(0) = x_0$, $p^{\pm}(0) = p_0$. Here,

$$H^{\pm}(x,p) = \pm c(x)|p| \tag{2.8}$$

are the signed Hamiltonians. For the sake of simplicity, we drop the superscript $\pm$. In particular, H denotes either H^+ or H^-.

Second, the matrix M satisfies the Riccati equation,

$$\dot{M}(t) + D(t) + B(t)M(t) + M(t)B(t)^t + M(t)X(t)M(t) = 0, \tag{2.9}$$

where $B(t), X(t), D(t)$ are $d \times d$ matrices with elements given by the second-order derivatives of the Hamiltonian,

$$D_{ij}(t) = \partial_{x_i}\partial_{x_j} H, \quad B_{ij}(t) = \partial_{x_i}\partial_{p_j} H, \quad X_{ij}(t) = \partial_{p_i}\partial_{p_j} H,$$

evaluated along the bicharacteristic $(x,p) = (x(t),p(t))$. The Riccati equation is supplemented with an initial condition $M(0) = M_0$. The symplectic structure of the Hamilton system implies that $M(t)$ is symmetric and has a positive definite imaginary part provided that it initially does (see Lemma 2.6 and [22, Lemma 2.56]).

Third, the amplitude function A satisfies the transport equation

$$\dot{A}(t) + \frac{A(t)}{2H}\Big(c^2\,\mathrm{Tr}(M(t)) - \partial_p H\cdot\partial_x H - (\partial_p H)^T M(t)\,\partial_p H\Big) = 0, \tag{2.10}$$

where H and its derivatives are evaluated along the bicharacteristic $(x,p) = (x(t),p(t))$. This equation is supplemented with an initial condition $A(0) = A_0$.

In what follows, we discuss families of Gaussian beams, with the goal of describing superimpositions and time evolution. See Figs. 2.3 and 2.4 for plots of a front of Gaussian beams going through a caustic.

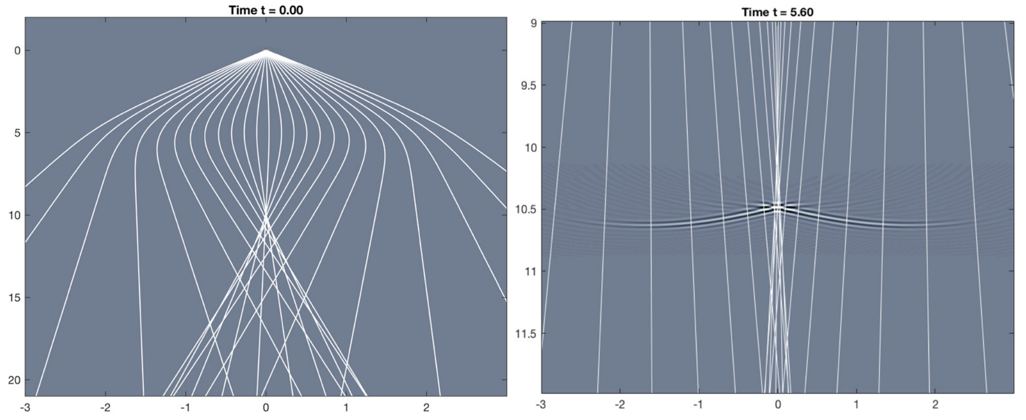

Fig. 2.3. The projected characteristics corresponding to the velocity $c(x_1, x_2) = 2 - 0.4 * \exp(-(x_1^2 + (x_2 - 5)^2)/3)$, and a detail of the evolution of a front of Gaussian packets, from $t = 0$ to $t = 8.40$. Initially localized on the boundary, the front goes through a caustic at time $t \approx 5.60$.

2.4. *Sets of initial conditions for Gaussian beams*

We consider families of Gaussian beams associated with sets of parameters described as follows. We let Γ_0 be a subset of Γ and $\mathcal{S}$ be a map

$$\mathcal{S}_\gamma = (\omega_\gamma, a_\gamma, \xi_\gamma, \mathcal{A}_\gamma, \mathcal{M}_\gamma) \in \mathbb{R}_+ \times \mathbb{R}^d \times (\mathbb{R}^d \setminus \{0\}) \times (\mathbb{R} \setminus \{0\}) \times \mathbb{C}^{d\times d}, \qquad \gamma \in \Gamma_0, \tag{2.11}$$

such that $\mathcal{M}_\gamma$ is symmetric and $\Im \mathcal{M}_\gamma > 0$ for all $\gamma \in \Gamma_0$. We associate two functions $\Phi_\gamma^+, \Phi_\gamma^- : \mathbb{R} \times \mathbb{R}^d \to \mathbb{C}$ that we now describe. To simplify the notation we drop the superscript $+, -$. Let $x_\gamma(t)$, $p_\gamma(t)$, $M_\gamma(t)$, $A_\gamma(t)$ be the solutions to the set of ODEs defined in (2.7), (2.9) and (2.10), supplemented with initial conditions:

$$\begin{cases} x|_{t=0} = a_\gamma, & p|_{t=0} = 2\pi \dfrac{\xi_\gamma}{\omega_\gamma}, \\ M|_{t=0} = 2\pi \mathcal{M}_\gamma, & A|_{t=0} = \mathcal{A}_\gamma \omega_\gamma^{\frac{d}{4}}. \end{cases} \tag{2.12}$$

We now define the beams by

$$\Phi_\gamma(t, x) = A_\gamma(t) e^{i\omega_\gamma \theta_\gamma(t,x)}, \tag{2.13}$$

with

$$\theta_\gamma(t, x) = p_\gamma(t) \cdot (x - x_\gamma(t)) + \tfrac{1}{2}(x - x_\gamma(t)) \cdot M_\gamma(t)(x - x_\gamma(t)). \tag{2.14}$$

The ODEs in (2.7), (2.9) and (2.10) have globally defined unique solutions for the initial conditions given by (2.12). Indeed, the system of ODEs in (2.7) is the flow associated with the Hamiltonian H and, due to the homogeneity of $H(x, p)$ in p, it is solvable as long as the initial condition $p|_{t=0}$ is non-zero. That is why we require that $\xi_\gamma \neq 0$. Once the Hamiltonian flow

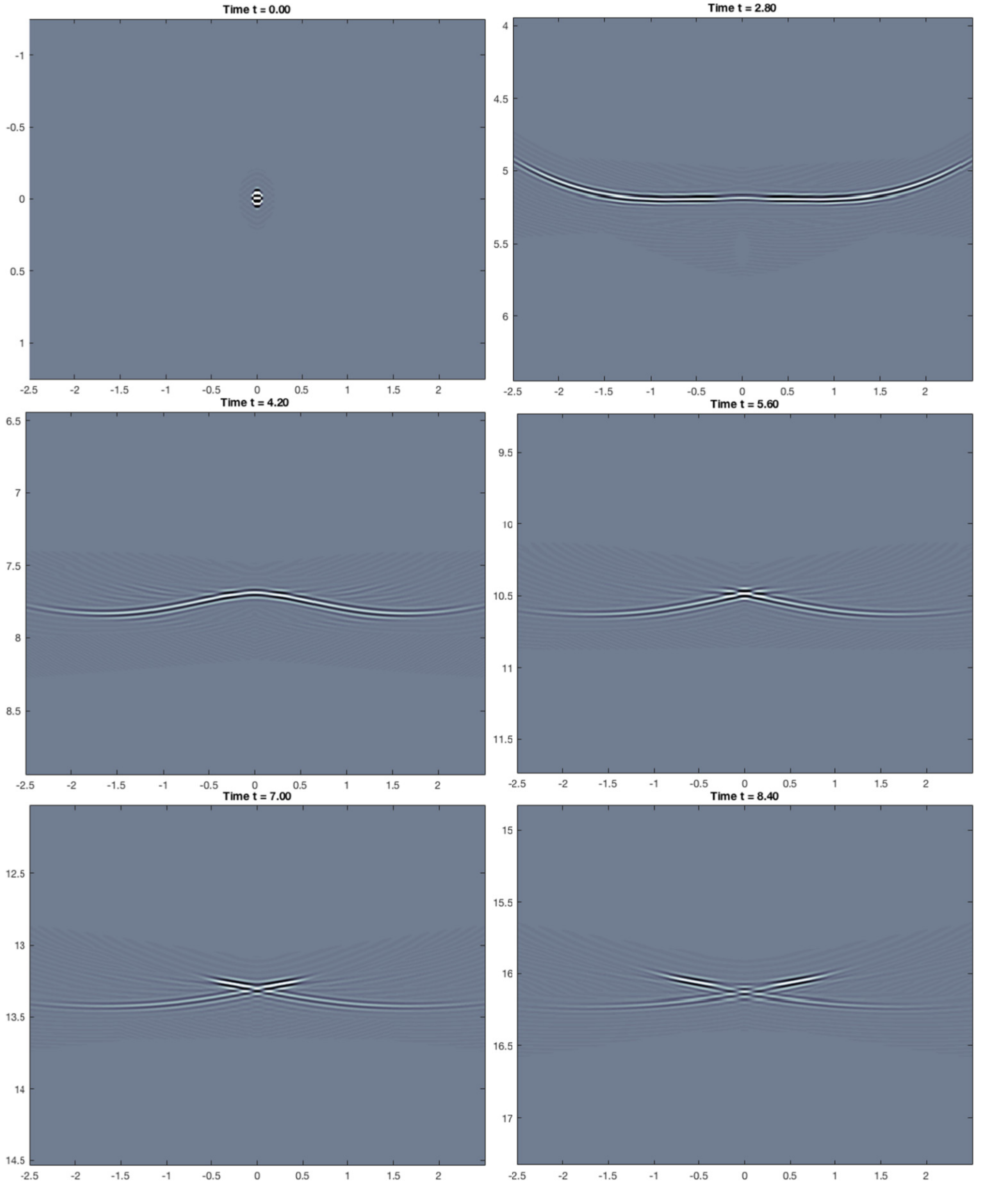

Fig. 2.4. Detailed evolution of the wavefront corresponding to Fig. 2.3. A tolerance of $\Im(M(t)) > 0.005$ was set, suppressing the beam if the condition was not satisfied.

(x, p) is defined, (2.9) has a globally defined unique solution because the initial datum is symmetric and has a positive imaginary part [22, Lemma 2.56]. Finally, (2.10) has also a unique global solution, since it is a linear ODEs with continuous coefficients.

Remark 2.3. When we need to emphasize the dependence on the choice of sign for H we write: $\Phi^{\pm}_{\gamma}$, $x^{\pm}_{\gamma}(t)$, $p^{\pm}_{\gamma}(t)$, $M^{\pm}_{\gamma}(t)$, $A^{\pm}_{\gamma}(t)$. We stress that these functions depend not only on the index

γ, but also on the underlying map from (2.11), that describes how to associate with γ initial conditions for the ODEs defining the beam. When we need to stress this dependence we use further superscripts.

Remark 2.4. By abuse of language, we often refer to a set of GB parameters $\Upsilon = \{\mathcal{S}_\gamma : \gamma \in \Gamma_0\}$, although it is not the set Υ, but the underlying map $\mathcal{S}$ that matters. Hence, Υ should be considered as an indexed set, that is formally equivalent to the map $\mathcal{S}$.

2.5. Standard initial conditions

We now describe the canonical set of GB parameters, defined so that the corresponding Gaussian beams at time $t = 0$ coincide with the higher-scale part of the frame $\mathcal{F}$. We define *the standard set of Gaussian beam parameters* as the set $\Upsilon^{st} = \left\{\mathcal{S}^{st}_\gamma : \gamma \in \Gamma\right\}$ given by

$$\mathcal{S}^{st}_\gamma = (4^j, 2^{-j}\lambda, \xi_{j,k}, 2^{\frac{d}{4}}, iI_d), \qquad \gamma = (j,k,\lambda) \in \Gamma. \tag{2.15}$$

The corresponding beams are denoted $\{\Phi^{\mathrm{st},\pm}_\gamma : \gamma \in \Gamma\}$.

Observation 2.5. *For the standard set of parameters* Υ^{st}*:*

$$\Phi^{\mathrm{st},\pm}_\gamma(0,x) = \varphi_{j,k,\lambda}(x), \qquad \gamma = (j,k,\lambda) \in \Gamma.$$

This follows by substituting (2.15) into (2.12) and (2.13)-(2.14).

2.6. Properties of the defining ODEs

We now show that certain uniformity properties of a family of Gaussian beams parameters imply corresponding uniformity properties for the ODEs defining the beams.

Lemma 2.6. *Let* $\{\mathcal{S}_\gamma : \gamma \in \Gamma_0\}$ *be a set of GB parameters. Assume that there exist* $0 < C_0 \le C_1$ *such that* $C_0 |\xi_\gamma| \le \omega_\gamma \le C_1 |\xi_\gamma|$*. Let* $T > 0$*. Then the following estimates hold for* $\gamma, \gamma' \in \Gamma_0$ *and* $t \in [-T, T]$*:*

$$|a_\gamma - a_{\gamma'}|^2 \le |x_\gamma(t) - x_{\gamma'}(t)|^2 + C_T,\ |x_\gamma(t) - x_{\gamma'}(t)|^2 \le |a_\gamma - a_{\gamma'}|^2 + C_T, \tag{2.16}$$

$$|p_\gamma(t)| \asymp |p_\gamma(0)| \asymp 1, \qquad |\dot{x}_\gamma(t)|, |\dot{p}_\gamma(t)| \lesssim 1, \tag{2.17}$$

$$d\Big(\big(x_\gamma(t), \omega_\gamma p_\gamma(t)\big), \big(x_{\gamma'}(t), \omega_{\gamma'} p_{\gamma'}(t)\big)\Big) \asymp d\Big(\big(a_\gamma, \xi_\gamma\big), \big(a_{\gamma'}, \xi_{\gamma'}\big)\Big), \tag{2.18}$$

where the constant C_T *and the implied constants depend on* T*,* C_0 *and* C_1 *but not on the particular pair of parameters* γ, γ'*.*

If, in addition, $\|\mathcal{M}_\gamma\| \le C_1$*, and* $\Im \mathcal{M}_\gamma \ge C_0 \cdot I_d$*, then*

$$\|M_\gamma(t)\| \lesssim 1, \qquad \Im M_\gamma(t) \gtrsim I_d, \qquad \left|\mathcal{A}_\gamma \omega_\gamma^{\frac{d}{4}}\right| \asymp \left|A_\gamma(t)\right|. \tag{2.19}$$

Proof. The bounds (2.16) and (2.17) follow from the assumptions on the velocity and Gronwall's lemma; see the proofs of [4, Lemmas 3.1 and 3.3]. Using the equations for the Hamiltonian flow, the assumptions on c, and (2.16) we get

$$\left|\dot{x}_\gamma(t)\right| = |c(x(t))| \leq C,$$
$$\left|\dot{p}_\gamma(t)\right| = |\nabla c(x(t))|\,|p(t)| \lesssim C,$$

where C is a constant that depends only on the velocity c. This gives (2.17). Finally, the estimate in (2.18) is proved in [4, Lemma 3.2].

For the "in addition" part, consider the matrix-valued ODE in (2.9). The derivatives of the Hamiltonian $H(x,p)$ are bounded on any set where $|p|$ is bounded above and below. Since $\left|p_\gamma(t)\right|$ is bounded above and below on $[-T,T]$ by Lemma 2.6, it follows that the coefficients in (2.9) are bounded. Second, the norm of the initial condition $M_\gamma(0) = \mathcal{M}_\gamma$ is bounded by assumption - cf. (2.12). Therefore, the first part of (2.19) follows by Gronwall's lemma. The second part in (2.19) now follows from [4, Lemma 3.1] (which requires $\|M(t)\|$ to be bounded). Finally, the third part of (2.19) is proved in [22, Lemma 2.56]. The statement there is non-quantitative, but the argument gives the desired conclusion. See also [4, Lemma 3.4]. □

Remark 2.7. In Lemma 2.6, the conclusion $\left|p_\gamma(t)\right| \asymp 1$ holds because the initial condition associated with γ in (2.12) ensures that $\left|p_\gamma(0)\right| \asymp 1$, with the assumption that $\left|\xi_\gamma\right| \asymp \omega_\gamma$. In general, if (x,p) is the flow associated with H^+ or H^- with arbitrary initial conditions, it follows from our assumptions in the velocity that $|p(t)| \asymp |p(0)|$ with constants that are uniform on any bounded interval of time.

3. Well-spread families of Gaussian beam parameters

3.1. Definitions

We develop criteria under which a family of Gaussian beam parameters behaves qualitatively like the standard one, given by

$$\mathcal{S}^{st}_\gamma = (\omega^{st}_\gamma, a^{st}_\gamma, \xi^{st}_\gamma, \mathcal{A}^{st}_\gamma, \mathcal{M}^{st}_\gamma) = (4^j, 2^{-j}\lambda, \xi_{j,k}, 2^{\frac{d}{4}}, iI_d), \qquad \gamma = (j,k,\lambda).$$

Our main goal is to show that when an adequate family of parameters is used as initial values, then a linear combination of the corresponding Gaussian beams satisfies a suitable Bessel bound and provides an approximate solution to the wave equation.

Definition 3.1. A *well-spread set of Gaussian beam parameters* is an indexed set

$$\Upsilon \equiv \left\{\mathcal{S}_\gamma : \gamma \in \Gamma_0\right\} \subseteq \mathbb{R}_+ \times \mathbb{R}^d \times (\mathbb{R}^d \setminus \{0\}) \times (\mathbb{R} \setminus \{0\}) \times \mathbb{C}^{d\times d}, \qquad \gamma \in \Gamma_0$$

with $\Gamma_0 \subseteq \Gamma$, such that

(i) $|a^{st}_\gamma - a^{st}_{\gamma'}| \lesssim |a_\gamma - a_{\gamma'}| + 1, \qquad \gamma, \gamma' \in \Gamma_0.$
(ii) $d((a_\gamma, \xi_\gamma), (a_{\gamma'}, \xi_{\gamma'})) \gtrsim d((a^{st}_\gamma, \xi^{st}_\gamma), (a^{st}_{\gamma'}, \xi^{st}_{\gamma'})), \qquad \gamma, \gamma' \in \Gamma_0.$
(iii) $\mathcal{M}_\gamma \in \mathbb{C}^{d\times d}$ is symmetric, $\|\mathcal{M}_\gamma\| \lesssim 1$ and $\Im(\mathcal{M}_\gamma) \gtrsim I_d, \qquad \gamma \in \Gamma_0.$

(iv) $\omega_\gamma \asymp 4^j$ and $|\xi_\gamma| \asymp \omega_\gamma$, $\quad \gamma \in \Gamma_0$.
(v) $|\mathcal{A}_\gamma| \asymp 1$, $\quad \gamma \in \Gamma_0$.

In the last definition, the symbols $\lesssim$, $\gtrsim$ and $\asymp$ should be interpreted as asserting the existence of suitable constants that are uniform within the family Υ.

Remark 3.2. For short, we say that $\{\Phi_\gamma^+ : \gamma \in \Gamma_0\}$ and $\{\Phi_\gamma^- : \gamma \in \Gamma_0\}$ are *well-spread families of Gaussian beams*, implying the existence of a corresponding well-spread family of Gaussian beam parameters $\Upsilon \equiv \{\mathcal{S}_\gamma : \gamma \in \Gamma_0\}$ that defines the beams.

Similarly, when a certain family of GB parameters Υ is discussed, we may denote the corresponding beams by just $\Phi_\gamma^\pm$, without remarking their dependence on the map $\mathcal{S}$.

Before proving the main estimates, we define an adequate notion of vanishing order along a family of Gaussian beams.

Definition 3.3. Given a well-spread set of GB parameters $\Upsilon \equiv \{\mathcal{S}_\gamma : \gamma \in \Gamma_0\}$, an interval $I \subseteq \mathbb{R}$, and $m \in \mathbb{N}_0$, a family of functions $F \equiv \{F_\gamma : \gamma \in \Gamma_0\}$ is said to be $F = \mathbb{O}^m(I, \Upsilon)$ if

- $F_\gamma(t,x) = \sum_{|\eta|=m} G_{\gamma,\eta}(t,x)(x - x_\gamma(t))^\eta$, for some functions $G_{\gamma,\eta}(t,\cdot) \in C_b^\infty(\mathbb{R}^d)$, for all $t \in I$.
- $\sup_{\gamma\in\Gamma_0, t\in I} \|\partial^k G_{\gamma,\eta}(t,\cdot)\|_{L^\infty(\mathbb{R}^d)} < +\infty$, for all multi-indices k and η, with $|\eta| = m$.

Thus, $F = \mathbb{O}^m(I, \Upsilon)$ means that each $F_\gamma(t,\cdot) = O(|x - x_\gamma(t)|^m, \mathbb{R}^d)$ and the corresponding bounds are uniform for $t \in I$ and $\gamma \in \Gamma_0$.

We note that the definition of $\mathbb{O}^m(I, \Upsilon)$ involves a vanishing condition at $x = x_\gamma(t)$ and also a growth condition for $|x - x_\gamma(t)| \gg 1$. As a consequence, $F = \mathbb{O}^{m+1}(I, \Upsilon)$ does not imply $F = \mathbb{O}^m(I, \Upsilon)$. As a remedy, we introduce the following notion.

Definition 3.4. Given a well-spread set of GB parameters $\Upsilon \equiv \{\mathcal{S}_\gamma : \gamma \in \Gamma_0\}$ and an interval $I \subseteq \mathbb{R}$, a family of functions $F \equiv \{F_\gamma : \gamma \in \Gamma_0\}$ is said to be $F = \mathbb{O}^m_\geq(I, \Upsilon)$ if there exists a finite family $F^1 = \mathbb{O}^{m_1}(I, \Upsilon), \ldots, F^n = \mathbb{O}^{m_n}(I, \Upsilon)$, with $m_1, \ldots, m_n \geq m$, such that $F_\gamma(t,x) = F_\gamma^1(t,x) + \ldots + F_\gamma^n(t,x)$.

Note that $F = \mathbb{O}^{m+1}_\geq(I, \Upsilon)$ implies that $F = \mathbb{O}^m_\geq(I, \Upsilon)$.

3.2. Bessel bounds and vanishing orders

The following Bessel bounds for the summation of Gaussian beams with factors vanishing at the spatial center of the beams extend those in [4, Sec. 3] from $L^2(\mathbb{R}^d)$ to Sobolev spaces, and to more general sets of initial conditions.

Theorem 3.5. *Let $\Upsilon = \{\mathcal{S}_\gamma : \gamma \in \Gamma_0\}$ be a well-spread set of Gaussian beam parameters and let $F = \mathbb{O}^m_\geq(I, \Upsilon)$, with $I \subseteq \mathbb{R}$ a bounded interval and $s \in [0,1]$. Then*

$$\sup_{t\in I} \Big\| \sum_{\gamma\in\Gamma_0} 2^{jm} b_\gamma \Phi_\gamma^\pm(t,\cdot) F_\gamma(t,\cdot) \Big\|_{H^s}^2 \lesssim C_I \sum_{\gamma\in\Gamma_0} 4^{2sj} |b_\gamma|^2,$$

with $b_\gamma \in \mathbb{C}$ such that the sum on the right-hand side is finite. (Here, the constant C_I depends on the interval I and the family F.)

A proof of Theorem 3.5 can be given by showing that for fixed time well-spread Gaussian beams are qualitatively similar to wave-atoms, and by studying the action of families of pseudodifferential operators on such systems. See [5, Appendix B and C] for full proofs.

Remark 3.6. The choice $I = \{0\}$ is allowed in Theorem 3.5 and corresponds to time-independent functions $F_\gamma(x)$ and frame elements: $\| \sum_{\gamma\in\Gamma} 2^{jm} b_\gamma \varphi_\gamma F_\gamma \|_{H^s}^2 \lesssim \sum_{\gamma\in\Gamma} 4^{2js} |b_\gamma|^2$.

3.3. Uniformity of errors for Taylor expansions

Most of our arguments rely on Taylor expansions for the functions x, p, A, M used in the definition of Gaussian beams. The following lemma is used to justify that, in such arguments, the error terms can be bounded uniformly within a given well-spread family of Gaussian beams.

Lemma 3.7. *Let $\Upsilon = \{\mathcal{S}_\gamma : \gamma \in \Gamma_0\}$ be a well-spread set of Gaussian beam parameters, let $T \geq 0$ and $k \geq 0$ be an integer. Then the following quantity:*

$$\sup_{\gamma\in\Gamma_0} \sup_{t\in[-T,T]} \left|\partial_t^{k+1} x_\gamma(t)\right| + \left|\partial_t^k p_\gamma(t)\right| + \left|\partial_t^k M_\gamma(t)\right|,$$

is bounded by a constant that depends on T, k and Υ. In addition,

$$\partial_t A_\gamma(t) = A_\gamma(t) G_\gamma(t), \tag{3.1}$$

with $\sup_{\gamma\in\Gamma_0} \sup_{t\in[-T,T]} \left|\partial_t^k G_\gamma\right|$ bounded by a constant that depends on T, k and Υ.

Proof. Since the derivatives of the velocity c are bounded, the derivatives of $H(x, p)$ are bounded on any set where $|p|$ is bounded above and below. By Lemma 2.6, $\left|p_\gamma(t)\right| \asymp 1$ and, therefore, we conclude that

$$\sup_{t\in[-T,T]} \left|\partial_x^n \partial_p^m H(x_\gamma(t), p_\gamma(t))\right| \lesssim C_{n,m,T} < \infty, \tag{3.2}$$

for all multi-indices n, m. Inspecting the definition of the Hamiltonian field (x_γ, p_γ) - cf. (2.7), the claim on x and p follows from (3.2).

For the matrix M_γ, we note that, due to (3.2), it satisfies a Riccati-type ODE where the coefficients are bounded and have all the derivatives bounded. Moreover, the corresponding initial condition is bounded, as part of Definition 3.1. Hence, the claim on M_γ follows from a Gronwall-type argument for linear systems of ODEs - see for example [9] and [22, Lemma 2.56].

Finally, inspecting (2.10), we see that the claim for the amplitude follows from (3.2) and the previous bounds. □

3.4. Asymptotic solutions

We now clarify how a linear combination of Gaussian beams with well-spread parameters approximately solves the wave equation. These results have been proved in [4] for standard Gaussian beam parameters, and here are extended to more general initial conditions.

Theorem 3.8. *Let $\Upsilon = \{\mathcal{S}_\gamma : \gamma \in \Gamma_0\}$ be a well-spread set of Gaussian beam parameters. Then*

$$\sup_{t\in[0,T]} \left\| \left(\partial_t^2 - c^2(x)\Delta_x\right) \sum_{\gamma\in\Gamma_0} b_\gamma \Phi_\gamma^\pm(t,\cdot) \right\|_{L^2(\mathbb{R}^d)}^2 \leq C_T \sum_{\gamma\in\Gamma_0} 4^j \left|b_\gamma\right|^2,$$

with $b_\gamma \in \mathbb{C}$ such that the sum on the right-hand side is finite.

The proof of Theorem 3.8 is based on the following description of the time derivatives of beams, which will be helpful later:

$$\partial_t \Phi_\gamma^\pm(t,x) = \left(F_\gamma^{(0)}(t,x) + 4^j F^{(1)}(t,x) - i\omega_\gamma H^\pm(x_\gamma^\pm(t), p_\gamma^\pm(t))\right)\Phi_\gamma^\pm(t,x), \tag{3.3}$$

with $F^{(m)} = \mathbb{O}_{\geq}^m([-T,T],\Upsilon)$. We omit the proofs and refer the interested reader to [5].

4. The Dirichlet problem on the half space

4.1. Setting and assumptions

We are interested in the following problem. Suppose that $u : [0,T]\times\mathbb{R}_+^d \to \mathbb{C}$ is a (weak) solution to:

$$\begin{cases} \partial_t^2 u(t,x) - c(x)^2\Delta_x u(t,x) = 0, & t\in[0,T], x\in\mathbb{R}_+^d, \\ u(0,x) = u_t(0,x) = 0, & x\in\mathbb{R}_+^d, \\ u(t,0,y) = h(t,y), & t\in[0,T], y\in\mathbb{R}^{d-1}, \end{cases}$$

where $h \in H^1([0,T]\times\mathbb{R}^{d-1})$ is called *boundary value*. We assume that we are able to measure the boundary value h and the goal is to approximate the corresponding solution u. We now introduce several assumptions.

4.1.1. Assumptions on the boundary value

In order for the Dirichlet problem to be well-posed we need to assume that h satisfies the standard compatibility condition $h(0,\cdot) \equiv 0$. In addition, the parametrix that we propose is ultimately based on oscillatory integrals and the theory of elliptic boundary value problems, and these techniques require that the bicharacteristic directions be nowhere tangent to the boundary [36]. That is why we exclude *grazing rays* from the wavefront set of h. Following [37], we formulate quantitative versions of these assumptions by replacing the function h with a new function h_{cut} that is the result of applying an adequate pseudodifferential cut-off to h. Recall that h is a function of $(t,x_*) \in \mathbb{R}\times\mathbb{R}^{d-1}$. We denote the conjugate (Fourier) variables by (τ,ξ_*), and let $h_{cut} := \eta(t,x_*,D_t,D_{x_*})h$, where the symbol $\eta(t,x_*,\tau,\xi_*) := a(t,x_*)b(t,x_*,\tau,\xi_*)$ satisfies the following.

(i) a is smooth with compact support and there exist $C_{h,\inf}, C_{h,\sup} \in (0,T)$ such that $\operatorname{supp}(a) \subseteq [C_{h,\inf}, C_{h,\sup}] \times \mathbb{R}^{d-1}$.
(ii) b is a smooth symbol of order 0, and there is a constant $C > 0$ such that b vanishes on the set of all points $(t, x_*, \tau, \xi_*) \in \mathbb{R} \times \mathbb{R}^{d-1} \times \mathbb{R} \times \mathbb{R}^{d-1}$ such that $|(\tau, \xi_*)| \geq 1$ and

$$\frac{|\tau|}{c(0,x_*)} - |\xi_*| \leq C\,|(\tau, \xi_*)|\,. \tag{4.2}$$

(This is possible because the condition in (4.2) is homogeneous of degree zero on (τ, ξ_*).) If such rays are not present in the wavefront set of the boundary value, the action of the cut-off is not needed.

We note that, as a result of the cut-off operation, $h_{cut} \in H^1([0,T] \times \mathbb{R}^{d-1})$, $\operatorname{supp}(h_{cut}) \subseteq [C_{h,\inf}, C_{h,\sup}] \times K$, for some compact set $K \subset \mathbb{R}^{d-1}$, and $[C_{h,\inf}, C_{h,\sup}] \subseteq (0,T)$.

Remark 4.1. The assumptions on the boundary value are quantitative versions of the compatibility and no-grazing ray conditions. Indeed, we assume that the observation window $[0,T]$ properly contains the time support of h, and that there is an absolute lower bound on the grazing angles.

4.1.2. Assumptions on the velocity

We recall that the *velocity* c is assumed to be smooth, positive, bounded from below and with bounded derivatives of all orders. This ensures that suitable energy estimates are available for the Dirichlet problem. *We note that although we aim to provide a parametrix for the right half-space, we assume that the velocity is defined on the whole Euclidean space*. This is just a matter of convenience, since, as a consequence of Theorem 7.1, the values of the velocity on the left half-space impact the parametrix only up to the parametrix error. Indeed, two different choices for c that agree on the right half-space lead to approximate solutions that are suitably close to the exact one, and thus close to each other.

4.1.3. The cone condition

We assume that for every $\varepsilon \in (0,1]$, there exist $\delta > 0$ such that if $(x(t), p(t))$ is a solution to the Hamiltonian flow with initial conditions at $t_0 \in [C_{h,\inf}, C_{h,\sup}]$ satisfying $x_1(t_0) = 0$ and $|p_1(t_0)| \geq \varepsilon |p(t_0)|$ then:

$$|x_1(t)| \geq \delta |t - t_0|, \qquad t \in [-T, T]. \tag{4.3}$$

Remark 4.2. The cone condition implies that for all take-off angles at the boundary the corresponding rays do not return to the boundary in the time interval in question, and indeed it is a quantitative version of that statement. See Fig. 4.5.

Remark 4.3. Since

$$|\dot{x}_1(t)| = c(x(t)) \frac{|p_1(t)|}{|p(t)|} \asymp \frac{|p_1(t)|}{|p(t)|},$$

the cone condition holds automatically for t near t_0. The content of (4.3) is the validity of the bound on the whole interval $[-T, T]$. Moreover, since the Hamiltonian is time independent, this

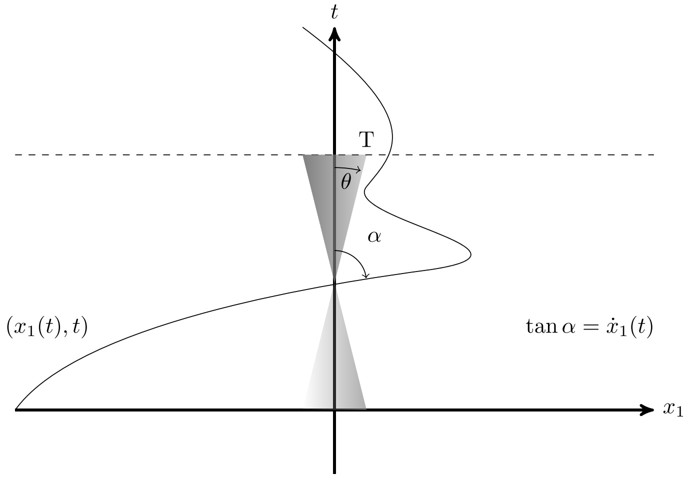

Fig. 4.5. The cone condition.

condition can be stated at $t_0 = 0$ and it is only about the size of the interval on which the cone condition holds. Fig. 2.3 shows an example of a velocity satisfying the hypothesis.

4.2. *Frame expansion of the boundary value*

The recovery method that we introduce in the next sections operates on the frame expansion of the boundary value

$$h(t, x_*) = \sum_{\gamma \in \Gamma_*} h_\gamma \varphi_\gamma(t, x_*).$$

Therefore, we need to show that the assumptions above are reflected by this expansion. Recall that $h_{cut} = \eta(t, x_*, D_t, D_{x_*})h$, where η is a zero-order pseudodifferential symbol. As shown in Section 2.2, this operator can be approximately implemented as a cut-off on the frame coefficients. More precisely, we first discard to zeroth-scale coefficients, then let $\tilde{h}_\gamma := \eta(2^{-j}\lambda, \xi_{j,k})h_\gamma$, $\Gamma_h := \{\gamma \in \Gamma : \tilde{h}_\gamma \neq 0\}$, and set

$$\tilde{h} = \sum_{\gamma \in \Gamma_h} \tilde{h}_\gamma \varphi_\gamma.$$

By Theorem 2.2, we have the following approximation estimate:

$$\|h_{cut} - \tilde{h}\|_{H^1} \lesssim \|h\|_{H^{1/2}}. \tag{4.4}$$

We now note some properties of the truncated frame parameters.

Proposition 4.4. *The set Γ_h satisfies the following.*

(i) *(*Time concentration and approximate compatibility*). There exist* $C_{h,\inf}, C_{h,\sup} > 0$ *such that for every* $\gamma = (j, k, \lambda) \in \Gamma_h$,

$$0 < C_{h,\inf} \leq 2^{-j}\lambda_1 \leq C_{h,\sup}. \tag{4.5}$$

(ii) *(*Quantitative grazing ray condition*). There exists* $C_{\text{graz}} \in (0, 1)$ *such that for every* $\gamma = (j, k, \lambda) \in \Gamma_h$*:*

$$\frac{|(\tilde{\xi}_{j,k})_1|}{c(0, 2^{-j}\lambda_*)} - |(\tilde{\xi}_{j,k})_*| \geq C_{\text{graz}}, \tag{4.6}$$

where the point $\tilde{\xi}_{j,k}$ *is defined by* (2.1)*.*

Proof. This follows directly from the properties of the symbol η. The constants $C_{h,\inf}, C_{h,\sup} > 0$ are the same as in Section 4.1.1. The constant C_{graz} is related to the constant C from (4.2). These two numbers are not exactly the same because the points $\tilde{\xi}_{j,k}$ are not exactly normalized - recall that, however, $\tilde{\xi}_{j,k}$ is a multiple of $\xi_{j,k}$ and $\left|\tilde{\xi}_{j,k}\right| \asymp 1$, so a suitable C_{graz} can be found. □

Remark 4.5. The constants $C_{h,\inf}, C_{h,\sup}, C_{\text{graz}}$ are given individual notation for future reference. We remark that the estimates in the rest of the article depend on them, as well as on the constants in the cone condition.

4.2.1. Non-tangential propagation

Since the velocity c is assumed to be bounded from below, the grazing ray condition (4.6) implies the following *non-tangential propagation* estimate:

$$\left|(\tilde{\xi}_{j,k})_1\right| \geq C, \qquad \gamma = (j, k, \lambda) \in \Gamma_h, \tag{4.7}$$

where $C = C_{\text{graz}} C_{\text{vel}} > 0$, and C_{vel} - cf. (1.10) - is the minimum value of the velocity c. In particular $(\tilde{\xi}_{j,k})_1 \neq 0$. In what follows, the sign of $(\tilde{\xi}_{j,k})_1$ plays an important role, and it is convenient to define:

$$\Gamma_h^+ := \left\{\gamma \in \Gamma_h : (\tilde{\xi}_{j,k})_1 < 0\right\}, \qquad \Gamma_h^- := \left\{\gamma \in \Gamma_h : (\tilde{\xi}_{j,k})_1 > 0\right\}. \tag{4.8}$$

(The motivation for this notation will be clear later.)

5. Spatio-temporal analysis of the beams near the boundary

We consider a well-spread family of Gaussian beams $\{\Phi_\gamma^+ : \gamma \in \Gamma_0\}$ or $\{\Phi_\gamma^- : \gamma \in \Gamma_0\}$, and times $t = t_\gamma$, $\gamma \in \Gamma_0$, at which the centers of the corresponding beams intersect the boundary $x_1 = 0$, i.e. $x_{\gamma,1}(t_\gamma) = 0$ - for short, we say that the *beams intersect the boundary* at those times. We focus on the case in which t_γ belongs to the interval $[C_{h,\inf}, C_{h,\sup}]$, where the boundary value is active. We assume that every beam in the family does intersect the boundary at a suitable time; for a more general family of beams, the analysis of this section applies by considering a subset of Γ_0.

We analyze the restriction of the beams to $x_1 = 0$, treating the remaining variables (t, x_*) as a *joint spatial variable*. We aim to approximately describe the restricted beam $\Phi^{\pm}_{\gamma}(t, 0, x_*)$ as a Gaussian beam with a fixed evolution time. We first identify the spatial center of $\Phi^{\pm}_{\gamma}(t, 0, x_*)$ and then describe the resulting functions in two different regimes: near the center and away from it. The assumption that the family of beams under study is well-spread allows us to obtain a uniform control on the approximation errors. This is essential for the applications in the following sections.

To ease the notation we focus on one of the two modes $(+/-)$ and remove this choice from the notation. Hence, most of the symbols below should be supplemented with a $+/-$ superscript. (In particular, H stands for either H^+ of H^-.)

5.1. Local analysis of a beam when it intersects the boundary

Before stating the estimates, we introduce some auxiliary functions defined in terms of the functions in (2.7), (2.9) and (2.10).

Let $\gamma \in \Gamma_0$ and consider the matrix $\widetilde{M}_{\gamma} \in \mathbb{C}^{d\times d}$ defined by

$$\begin{cases} \widetilde{M}_{\gamma,11} &= \dot{x}_{\gamma}(t_{\gamma}) \cdot M_{\gamma}(t_{\gamma})\dot{x}_{\gamma}(t_{\gamma}) - \dot{p}_{\gamma}(t_{\gamma}) \cdot \dot{x}_{\gamma}(t_{\gamma}), \\ \widetilde{M}_{\gamma,1k} &= \dot{p}_{\gamma,k}(t_{\gamma}) - \sum_{n=1}^{d} \big(M_{\gamma}(t_{\gamma})\big)_{kn} \dot{x}_{\gamma,n}(t_{\gamma}), \qquad k = 2, \dots, d, \\ \widetilde{M}_{\gamma,kl} &= \big(M_{\gamma}(t_{\gamma})\big)_{kl}, \qquad k, l = 2, \dots, d. \end{cases}$$

In more compact notation,

$$\widetilde{M}_{\gamma} = \begin{bmatrix} \widetilde{M}_{\gamma,11} & \widetilde{M}^{t}_{\gamma,1*} \\ \widetilde{M}_{\gamma,1*} & \big(M_{\gamma}(t_{\gamma})\big)_{**} \end{bmatrix},$$

where

$$\widetilde{M}_{\gamma,1*} = \big(\dot{p}_{\gamma}(t_{\gamma}) - M_{\gamma}(t_{\gamma})\dot{x}_{\gamma}(t_{\gamma})\big)_{*} \in \mathbb{C}^{(d-1)\times 1},$$

and $\big(M(t_{\gamma})\big)_{**} \in \mathbb{C}^{(d-1)\times(d-1)}$ is the matrix obtained from $M(t_{\gamma})$ by eliminating the first row and column. Let us also consider the following constants and functions:

$$\tau_{\gamma} = -H(x_{\gamma}(t_{\gamma}), p_{\gamma}(t_{\gamma})) = -p_{\gamma}(t_{\gamma}) \cdot \dot{x}_{\gamma}(t_{\gamma}), \tag{5.1}$$

$$L_{\gamma}(t, x_*) = \big(\tau_{\gamma}, p_{\gamma,*}(t_{\gamma})\big) \cdot ((t, x_*) - (t_{\gamma}, x_{\gamma,*}(t_{\gamma}))), \tag{5.2}$$

$$Q_{\gamma}(t, x_*) = \tfrac{1}{2}\big((t, x_*) - (t_{\gamma}, x_{\gamma,*})\big) \cdot \widetilde{M}(t_{\gamma})\big((t, x_*) - (t_{\gamma}, x_{\gamma,*})\big). \tag{5.3}$$

We can now describe a Gaussian beam intersecting the boundary.

Lemma 5.1. *Let $\Upsilon \equiv \{\mathcal{S}_{\gamma} : \gamma \in \Gamma_0\}$ be a well-spread set of GB parameters. For $\gamma \in \Gamma_0$, let $t_{\gamma} \in [C_{h,\inf}, C_{h,\sup}]$ be such that $x_{\gamma,1}(t_{\gamma}) = 0$ (i.e. the center of the corresponding beam $\Phi_{\gamma} = \Phi^{\pm}_{\gamma}$ intersects the boundary $x_1 = 0$ at a time $t = t_{\gamma}$ when the boundary value is active). Let us write*

$x_\gamma(t_\gamma) = (0, x_{\gamma,*}(t_\gamma))$. *Then the restriction of* Φ_γ *to* $x_1 = 0$ *admits the following asymptotic expansion around* $(t_\gamma, x_{\gamma,*}(t_\gamma))$*:*

$$\Phi_\gamma(t, 0, x_*) = A_\gamma(t_\gamma)\left(1 + R_\gamma(t)\right) e^{i\omega_\gamma\{L_\gamma(t,x_*)+Q_\gamma(t,x_*)+\Theta_\gamma(t,x_*)\}}, \ (t, x_*) \in \mathbb{R}^d_T, \tag{5.4}$$

with

$$R_\gamma(t) = r_\gamma(t)(t - t_\gamma), \quad \Theta_\gamma(t, x_*) = \sum_{|\mu|=3} g_{\gamma,\mu}(t)\left((t, x_*) - (t_\gamma, x_{\gamma,*}(t_\gamma))\right)^\mu$$

and $r_\gamma, g_{\gamma,\mu} \in C_b^\infty([-T, T])$*, uniformly on* γ*. More precisely, for every* $k \geq 0$*, the error factors satisfy:*

$$\sup_{\gamma\in\Gamma_0} \sup_{t\in[-T,T]} \left|\partial_t^k g_{\gamma,\mu}(t)\right|, \quad \sup_{\gamma\in\Gamma_0} \sup_{t\in[-T,T]} \left|\partial_t^k r_\gamma(t)\right| < +\infty. \tag{5.5}$$

Proof. We analyze the Gaussian beam

$$\Phi_\gamma(t, x) = A_\gamma(t)e^{i\omega_\gamma\theta_\gamma(t,x)},$$

by Taylor expanding the amplitude and phase.

Step 1. *The amplitude*. Using the bounds in Lemma 3.7 - specifically (3.1) - and Lemma 2.6 - which is applicable uniformly for $\gamma \in \Gamma_0$ - we see that the function $B_\gamma(t) := A_\gamma(t)/A_\gamma(t_\gamma)$ is bounded and has bounded derivatives on $[-T, T]$, uniformly for $\gamma \in \Gamma_0$. Since $B_\gamma(t_\gamma) = 1$, we can write: $B_\gamma(t) = 1 + r_\gamma(t)(t - t_\gamma)$, with r_γ as in (5.5). Therefore,

$$A_\gamma(t) = A_\gamma(t_\gamma)(1 + r_\gamma(t))(t - t_\gamma).$$

In order to establish (5.4), it remains to inspect the exponential factor.

Step 2. *Expansion of the characteristic flow*. We first expand the characteristics as

$$x_\gamma(t) = x_\gamma(t_\gamma) + \dot{x}_\gamma(t_\gamma)(t - t_\gamma) + \tfrac{1}{2}\ddot{x}_\gamma(t_\gamma)(t - t_\gamma)^2 + R_{x,\gamma}(t)(t - t_\gamma)^3, \tag{5.6}$$

$$p_\gamma(t) = p_\gamma(t_\gamma) + \dot{p}_\gamma(t_\gamma)(t - t_\gamma) + R_{p,\gamma}(t)(t - t_\gamma)^2, \tag{5.7}$$

where $R_{x,\gamma}, R_{p,\gamma} \in C_b^\infty([-T, T])$, and the corresponding bounds are uniform for $\gamma \in \Gamma_0$, as shown in Lemma 3.7.

We now focus on the phase function

$$\theta_\gamma(x, t) = p_\gamma(t) \cdot (x - x_\gamma(t)) + \tfrac{1}{2}(x - x_\gamma(t)) \cdot M_\gamma(t)(x - x_\gamma(t)).$$

Step 3. *The linear part of the phase*. The linear part of θ_γ is

$$\begin{aligned} p_\gamma(t) \cdot \left(x - x_\gamma(t)\right) &= p_\gamma(t_\gamma) \cdot \left(x - x_\gamma(t_\gamma) - \dot{x}_\gamma(t_\gamma)(t - t_\gamma) - \tfrac{1}{2}\ddot{x}_\gamma(t_\gamma)(t - t_\gamma)^2\right) \\ &\quad + \dot{p}_\gamma(t_\gamma)(t - t_\gamma) \cdot \left(x - x_\gamma(t_\gamma) - \dot{x}_\gamma(t - t_\gamma)\right) + \Theta_\gamma \\ &= p_\gamma(t_\gamma) \cdot \left(x - x_\gamma(t_\gamma)\right) - p_\gamma(t_\gamma)\dot{x}_\gamma(t_\gamma)(t - t_\gamma) \end{aligned} \tag{5.8}$$

$$
\begin{aligned}
&- \tfrac{1}{2}(t-t_\gamma)^2 \left(p_\gamma(t_\gamma)\cdot \ddot{x}_\gamma(t_\gamma) + \dot{p}_\gamma(t_\gamma)\cdot \dot{x}_\gamma(t_\gamma)\right) \qquad (5.9)\\
&- \tfrac{1}{2}(t-t_\gamma)^2 \left(\dot{p}_\gamma(t_\gamma)\cdot \dot{x}_\gamma(t_\gamma)\right)\\
&+ (t-t_\gamma)\dot{p}_\gamma(t_\gamma)\cdot \left(x - x_\gamma(t_\gamma)\right) + \Theta_\gamma,
\end{aligned}
$$

where Θ_γ denotes a function of the form:

$$
\Theta_\gamma = \sum_{|\mu|=3} R_{(x,p),\gamma,\mu}(t)\left((t,x)-(t_\gamma, x_\gamma(t_\gamma))\right)^\mu, \qquad R_{(x,p),\gamma,\mu} \in C_b^\infty([-T,T]).
$$

Indeed, note that the error factors $R_{(x,p),\gamma,\mu}(t)$ involve the error factors $R_{x,\gamma}$, $R_{p,\gamma}$ from (5.6) and (5.7) multiplied by $\dot{x}_\gamma(t_\gamma)$, $p_\gamma(t_\gamma)$ and similar quantities involving higher order derivatives, which are uniformly bounded by Lemma 3.7.

Since

$$
H(x_\gamma(t), p_\gamma(t)) = p_\gamma(t)\cdot \dot{x}_\gamma(t)
$$

is constant on t, it follows that

$$
\left(p_\gamma(t_\gamma)\cdot \ddot{x}_\gamma(t_\gamma) + \dot{p}_\gamma(t_\gamma)\cdot \dot{x}_\gamma(t_\gamma)\right) = \partial_t H(x_\gamma(t), p_\gamma(t))_{|t=t_\gamma} = 0,
$$

and the term in (5.9) vanishes. Thus, (5.8) reads

$$
\begin{aligned}
p_\gamma(t)\cdot\left(x - x_\gamma(t)\right) &= p_\gamma(t_\gamma)\cdot\left(x - x_\gamma(t_\gamma)\right) - p_\gamma(t_\gamma)\cdot \dot{x}_\gamma(t_\gamma)(t-t_\gamma)\\
&\quad - \tfrac{1}{2}(t-t_\gamma)^2\left(\dot{p}_\gamma(t_\gamma)\cdot \dot{x}_\gamma(t_\gamma)\right) + (t-t_\gamma)\dot{p}_\gamma(t_\gamma)\cdot\left(x - x_\gamma(t_\gamma)\right) + \Theta_\gamma.
\end{aligned}
$$

Specializing on the boundary we obtain that for $x_1 = x_{\gamma,1}(t_\gamma) = 0$,

$$
\begin{aligned}
p_\gamma(t)\cdot\left(x - x_\gamma(t)\right) &= L_\gamma(t, x_*) - \tfrac{1}{2}\left(\dot{p}_\gamma(t_\gamma)\cdot \dot{x}_\gamma(t_\gamma)\right)(t-t_\gamma)^2\\
&\quad + (t-t_\gamma)\dot{p}_{\gamma,*}(t_\gamma)\cdot\left(x_* - x_{\gamma,*}(t_\gamma)\right) + \Theta_\gamma|_{x_1=0},
\end{aligned}
\qquad (5.10)
$$

where $L(t, x_*)$ is defined by (5.2).

Step 4. *The quadratic part of the phase*. We linearize the Riccati matrix M_γ as

$$
M_\gamma(t) = M_\gamma(t_\gamma) + N_\gamma(t)(t-t_\gamma),
$$

with $N_\gamma \in C_b^\infty([-T,T])$ uniformly on γ, due to Lemma 3.7.

Using (5.6), we can expand the quadratic part of θ_γ as

$$
\begin{aligned}
&\left(x - x_\gamma(t)\right)\cdot M_\gamma(t)\left(x - x_\gamma(t)\right)\\
&= \left(x - x_\gamma(t)\right)\cdot M_\gamma(t_\gamma)\left(x - x_\gamma(t)\right) + \Theta_\gamma\\
&= \left(x - x_\gamma(t_\gamma) - \dot{x}_\gamma(t_\gamma)(t-t_\gamma)\right)\cdot M_\gamma(t_\gamma)\left(x - x_\gamma(t_\gamma) - \dot{x}_\gamma(t_\gamma)(t-t_\gamma)\right) + \Theta_\gamma,
\end{aligned}
$$

where, in each line, Θ_γ denotes a function of the form:

$$\Theta_\gamma = \sum_{|\mu|=3} R_{(x,M),\gamma,\mu}(t)\left((t,x)-(t_\gamma,x_\gamma(t_\gamma))\right)^\mu, \qquad \text{with } R_{(x,M),\gamma,\mu} \in C_b^\infty([-T,T]),$$

uniformly on γ.

Specializing on the boundary we obtain that for $x_1 = x_{\gamma,1}(t_\gamma) = 0$,

$$\begin{aligned}\left(x - x_\gamma(t)\right)\cdot M_\gamma(t)\left(x-x_\gamma(t)\right) &= \dot{x}_\gamma(t_\gamma)\cdot M_\gamma(t_\gamma)\dot{x}_\gamma(t_\gamma)(t-t_\gamma)^2 \\ &+ \left(x_* - x_{\gamma,*}(t_\gamma)\right)\cdot M_\gamma(t_\gamma)_{**}\left(x_*-x_{\gamma,*}(t_\gamma)\right) \\ &- 2\left(\dot{x}_\gamma(t_\gamma)\cdot M_\gamma(t_\gamma)\right)_*\left(x_*-x_{\gamma,*}(t_\gamma)\right)(t-t_\gamma) + \Theta_\gamma|_{x_1=0}.\end{aligned} \tag{5.11}$$

Step 5. *Collecting terms*. Finally, we combine (5.10) and (5.11), noting that the quadratic terms add up precisely to $Q_\gamma(t,x_*)$, as defined by (5.3). □

5.2. Global analysis of a beam when it intersects the boundary

We now describe the global profile of the restriction of the beam to the boundary $\{x_1 = 0\}$. We aim to show that the restricted beams $\Phi_\gamma(t,0,x_*)$ display a Gaussian profile *in the* (t,x_*) *variables*. To this end, we consider an additional assumption on the way in which the original beams intersect the boundary. We say that a family of beams $\{\Phi_\gamma : \gamma \in \Gamma_0\}$ *intersects the boundary in a uniformly transversal fashion* at times $\{t_\gamma : \gamma \in \Gamma_0\}$ if: (i) $x_{\gamma,1}(t_\gamma) = 0$, for all $\gamma \in \Gamma_0$, and (ii) there exists a constant $C_1 \in (0,1)$ such that

$$\left|p_{\gamma,1}(t_\gamma)\right| \geq C_1\left|p_\gamma(t_\gamma)\right|, \qquad \gamma \in \Gamma_0. \tag{5.12}$$

The following lemma provides the desired description.

Lemma 5.2. *Let $\Upsilon \equiv \left\{\mathcal{S}_\gamma : \gamma \in \Gamma_0\right\}$ be a well-spread set of GB parameters. For $\gamma \in \Gamma_0$, let $t_\gamma \in [C_{h,\inf}, C_{h,\sup}]$ be such that $x_{\gamma,1}(t_\gamma) = 0$ (i.e. the center of the corresponding beam $\Phi_\gamma = \Phi_\gamma^\pm$ intersects the boundary $x_1 = 0$ at a time $t = t_\gamma$ when the boundary value is active). Assume also that the beams intersect the boundary in a uniformly transversal fashion; i.e., there exists a constant $C_1 \in (0,1)$ such that (5.12) holds.*

Let us write $x_\gamma(t_\gamma) = (0, x_{\gamma,*}(t_\gamma))$. Then the restriction of Φ_γ to $x_1 = 0$ admits the following description: for $\gamma = (j,k,\lambda) \in \Gamma_0$,

$$\Phi_\gamma(t,0,x_*) = A_\gamma(t_\gamma)\exp\left(i4^j\left[L_\gamma(t,x_*) + i\ell\left((t-t_\gamma)^2 + \left|x_*-x_{\gamma,*}(t_\gamma)\right|^2\right)\right]\right)\cdot R_\gamma(t,x_*),$$

where L_γ is given by (5.2), $\ell > 0$ is a constant - that depends only on the family Υ and the constant C_1 - and $R_\gamma \in C_b^\infty([-T,T]\times\{x_* : \left|x_*-x_{\gamma,*}\right| \geq 1\})$, uniformly on γ. More precisely, for all multi-indices k, α, the error factor satisfies:

$$\sup_{\gamma\in\Gamma_0}\sup_{t\in[-T,T]}\sup_{\left|x_*-x_{\gamma,*}\right|\geq 1}\left|\partial_t^k\partial_{x_*}^\alpha R_\gamma(t,x_*)\right| < +\infty.$$

Proof. As before, all estimates in this proof are to be understood as being uniform for $\gamma = (j,k,\lambda) \in \Gamma_0$, and to be dependent on T.

Step 1. *Linearization of the centers*. Using Lemma 3.7 we write

$$x_\gamma(t) = x_\gamma(t_\gamma) + (t - t_\gamma)y_\gamma(t),$$

where $y_{\gamma,i} \in C_b^\infty([-T,T])$. Since $x_{\gamma,i}(t_\gamma) = 0$, the transversality assumption and the *cone condition* in (4.3), imply that

$$\left|y_{\gamma,1}(t)\right| \geq \delta, \qquad t \in [-T,T], \tag{5.13}$$

for some constant $\delta > 0$.

Step 2. *The linear part of the phase*. We show that

$$p_\gamma(t) \cdot \big((0,x_*) - x_\gamma(t)\big) = L_\gamma(t,x_*) + E_\gamma^1(t,x_*) \tag{5.14}$$

where E_γ^1 satisfies the following: given multi-indices k, α:

$$\sup_{t\in[-T,T]} \left|\partial_t^k \partial_{x_*}^\alpha E^1(t,x_*)\right| \leq C_{k,m}\left(1 + \left|x_* - x_{\gamma,*}(t_\gamma)\right|\right). \tag{5.15}$$

(Recall that this estimate is understood to be also uniform on γ, but dependent on T.)

We expand the left-hand side of (5.14). We use E to denote a function satisfying a bound similar to (5.15). The meaning of E changes from line to line, and the assertions are verified using Step 1, Lemma 2.6 and 3.7. With this understanding:

$$\begin{aligned}
p_\gamma(t) \cdot \big((0,x_*) - x_\gamma(t)\big) &= p_\gamma(t) \cdot \big((0,x_*) - x_\gamma(t_\gamma)\big) + E(t,x_*) \\
&= p_\gamma(t_\gamma) \cdot \big((0,x_*) - x_\gamma(t_\gamma)\big) + E(t,x_*) \\
&= p_{\gamma,*}(t_\gamma) \cdot \big(x_* - x_{\gamma,*}(t_\gamma)\big) + E(t,x_*) \\
&= L_\gamma(t,x_*) - \tau_\gamma(t - t_\gamma) + E(t,x_*) \\
&= L_\gamma(t,x_*) + E(t,x_*),
\end{aligned}$$

as desired.

Step 3. *The quadratic part of the phase*. Consider the quadratic term:

$$\mathcal{Q}(t,x_*) = \tfrac{1}{2}\left[((0,x_*) - x_\gamma(t)) \cdot M_\gamma(t)((0,x_*) - x_\gamma(t))\right].$$

Let us show that $\mathcal{Q}(t,x_*) = \mathcal{Q}^1(t,x_*) + \mathcal{Q}^2(t,x_*) + \mathcal{Q}^3(t,x_*)$, with

$$\mathcal{Q}^1(t,x_*) = \ell\left((t-t_\gamma)^2 + \left|x_* - x_{\gamma,*}(t_\gamma)\right|^2\right), \tag{5.16}$$

$$\mathcal{Q}^2(t,x_*) = \big(t - t_\gamma, x_* - x_{\gamma,*}(t_\gamma)\big) \cdot N_\gamma^2(t)\left(t - t_\gamma, x_* - x_{\gamma,*}(t_\gamma)\right), \tag{5.17}$$

$$\mathcal{Q}^3(t,x_*) = \big((0,x_*) - x_\gamma(t)\big) \cdot N_\gamma^3(t)\left((0,x_*) - x_\gamma(t)\right),$$

where $\ell > 0$, $N_\gamma^2(t), N_\gamma^3(t) \in \mathbb{C}^{d\times d}$ are symmetric, $\Im N_\gamma^2(t), \Im N_\gamma^3(t) \geq \ell' I_d$, $\ell' > 0$, and for each $k \geq 0$, there is a constant C_k such that

$$\sup_{t\in[-T,T]}\left|\partial_t^k N_\gamma^2(t)\right|, \sup_{t\in[-T,T]}\left|\partial_t^k N_\gamma^3(t)\right| \le C_k < +\infty. \tag{5.18}$$

By definition of well-spread set of GB parameters and Lemma 3.7, there exists a constant $\varepsilon > 0$ (independent of γ and t) such that $\Im(M_\gamma(t)) \ge 2\varepsilon I_d$. We let $N_\gamma^3(t) = \frac{1}{2}M_\gamma(t) - \frac{\varepsilon}{2} i I_d$. This defines $\mathcal{Q}^3$. Note that $\Im(N_\gamma^3(t)) \ge \ell' I_d$, with $\ell' = \frac{\varepsilon}{2}$ and that $\mathcal{Q}(t,x_*) - \mathcal{Q}^3(t,x_*) = \frac{\varepsilon}{2} i \left|(0,x_*) - x_\gamma(t)\right|^2$. Expanding that expression and using $x_{\gamma,1}(t_\gamma) = 0$, we see that

$$\begin{aligned}
\tfrac{2}{\varepsilon i}(\mathcal{Q}(t,x_*) - \mathcal{Q}^3(t,x_*)) &= \left|(0,x_*) - x_\gamma(t_\gamma) - (t-t_\gamma)y_\gamma(t)\right|^2 \\
&= \left|y_\gamma(t)\right|^2 (t-t_\gamma)^2 - 2(t-t_\gamma)y_{\gamma,*}(t)\cdot(x_* - x_{\gamma,*}(t_\gamma)) + \left|x_* - x_{\gamma,*}(t_\gamma)\right|^2 \\
&= (t-t_\gamma, x_* - x_{\gamma,*}(t_\gamma)) \cdot \widetilde{N}_\gamma(t)(t-t_\gamma, x_* - x_{\gamma,*}(t_\gamma)),
\end{aligned}$$

where:

$$\begin{bmatrix} \widetilde{N}_{\gamma,11}(t) & \widetilde{N}_{\gamma,1*}(t)^t \\ \widetilde{N}_{\gamma,1*}(t) & \widetilde{N}_{\gamma,**}(t) \end{bmatrix} = \begin{bmatrix} \left|y_\gamma(t)\right|^2 & -y_{\gamma,*}(t)^t \\ -y_{\gamma,*}(t) & I_{d-1} \end{bmatrix}.$$

Since $\left|y_\gamma(t)\right|^2 - \left|y_{\gamma,*}(t)\right|^2 = \left|y_{\gamma,1}(t)\right|^2$ is bounded below by (5.13), elementary linear algebra now shows that $\widetilde{N_\gamma}(t) \gtrsim I_d$, for $t \in [-T,T]$ - see [5, Lemma A.1] for details.

Hence, we can let $N^2(t) = \frac{\varepsilon}{2} i \widetilde{N}_\gamma(t) - \ell i I_d$ with $\ell > 0$ such that $\Im N^2(t) \ge \ell I_d$. We now let $\mathcal{Q}^1(t,x_*)$ and $\mathcal{Q}^2(t,x_*)$ be defined by (5.16) and (5.17), respectively. Hence, $\mathcal{Q}(t,x_*) = \mathcal{Q}^1(t,x_*) + \mathcal{Q}^2(t,x_*) + \mathcal{Q}^3(t,x_*)$ as desired. Finally the bounds in (5.18) follow from Lemma 3.7.

Step 4. *Bounds for the error factor*. We write $R(t,x_*) = R^1(t,x_*) \cdot R^2(t,x_*)$, with

$$\begin{aligned}
R^1(t,x_*) &= \frac{A(t)}{A(t_\gamma)}, \\
R^2(t,x_*) &= \exp(i4^j(E^1(t,x_*) + \mathcal{Q}^2(t,x_*) + \mathcal{Q}^3(t,x_*)))
\end{aligned}$$

By Lemmas 2.6, and 3.7, it follows that $R^1 \in C_b^\infty(\mathbb{R}_T^d)$ - cf. Step 1 in the proof of Lemma 5.1. We focus now on R^2. Let k, m be multi-indices. Using the bounds in Steps 2 and 3 (and the fact that E_γ^1 is real) we conclude that there exists a number $n = n(k,m)$ and a constant $C_n = C_{k,m}$ such that for $(t,x_*) \in \mathbb{R}^d{}_T$:

$$\begin{aligned}
\left|\partial_t^k \partial_{x_*}^m R^2(t,x_*)\right| \le C_n 4^{jn} \left(1 + \left|x_* - x_{\gamma,*}(t_\gamma)\right|^2\right)^n \cdot \\
\exp(-4^j \left[\Im(\mathcal{Q}^2(t,x_*)) + \Im(\mathcal{Q}^3(t,x_*))\right]).
\end{aligned} \tag{5.19}$$

Using (5.15) and the fact that $\Im N_\gamma^2, \Im N_\gamma^3 \ge \ell' I_d$ we obtain:

$$\begin{aligned}
\Im(\mathcal{Q}^2(t,x_*)) + \Im(\mathcal{Q}^3(t,x_*)) &\ge \ell' \left(\left|(t-t_\gamma)\right|^2 + \left|x_* - x_{\gamma,*}(t_\gamma)\right|^2 + \left|(0,x_*) - x_\gamma(t)\right|^2\right) \\
&\ge \ell' \left|x_* - x_{\gamma,*}(t_\gamma)\right|^2.
\end{aligned}$$

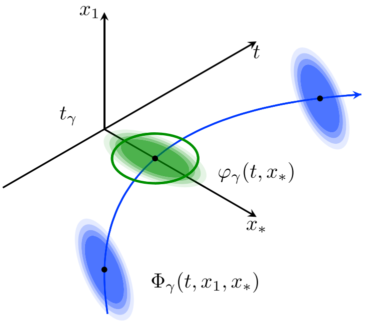

Fig. 6.6. A beam intersects the boundary $x_1 = 0$ at $t = t_\gamma$ moving along the projected bicharacteristic $x_\gamma(t)$. The frame element - represented by a circle - is matched to a beam - represented by a filled ellipse - which has an approximate Gaussian profile in (t, x_*).

Combining this with (5.19) we obtain:

$$\left|\partial_t^k \partial_{x_*}^m R^2(t, x_*)\right| \lesssim 4^{jn} \left(1 + \left|x_* - x_{\gamma,*}(t_\gamma)\right|^2\right)^n \exp(-4^j \ell' \left|x_* - x_{\gamma,*}(t_\gamma)\right|^2),$$

where the implied constant depends on k and m. Finally, for $\left|x_* - x_{\gamma,*}(t_\gamma)\right| \geq 1$ we can estimate:

$$\begin{aligned}\left|\partial_t^k \partial_{x_*}^m R^2(t, x_*)\right| &\lesssim 4^{jn} \left|x_* - x_{\gamma,*}(t_\gamma)\right|^{2n} \exp(-4^j \ell' \left|x_* - x_{\gamma,*}(t_\gamma)\right|^2) \\ &\lesssim \left(4^j \ell' \left|x_* - x_{\gamma,*}(t_\gamma)\right|^2\right)^n \exp(-4^j \ell' \left|x_* - x_{\gamma,*}(t_\gamma)\right|^2) \leq n!\end{aligned}$$

This completes the proof. □

6. Packet-beam matching

The goal of this section is to select, for each index $\gamma = (j, k, \lambda) \in \Gamma_h$, a corresponding tuple of initial conditions $\mathcal{S}_\gamma^h \in \mathbb{R}_+ \times \mathbb{R}^d \times (\mathbb{R}^d \setminus \{0\}) \times (\mathbb{R} \setminus \{0\}) \times \mathbb{C}^{d\times d}$ and an adequate mode, $+$ *or* $-$, giving initial conditions for a Gaussian beam, in such a way that

$$\Phi_\gamma^h(t, 0, x_*) \approx \varphi_\gamma(t, x_*). \tag{6.1}$$

We use the analysis of Section 5 as a guide. We first judiciously select a time instant t_γ and construct the beam Φ_γ^h in such a way that it intersects the boundary $\{x_1 = 0\}$ at that time. To this end, we design the beam Φ_γ^h by matching the approximate description of $\Phi_\gamma^h(t, 0, x_*)$, provided by Lemma 5.1, to the target frame element $\varphi_\gamma(t, x_*)$. This approximate description is useful for t near the boundary meeting time t_γ. Hence, the construction involves *back-propagating* the profile of the beam under construction by means of the ODEs in (2.7), (2.9), (2.10), from time $t = t_\gamma$ to time $t = 0$. The matching procedure is depicted in Fig. 6.6.

Afterwards, we analyze the family of parameters that results from this procedure, and prove that they are well-spread in the sense of Section 3, and that they intersect the boundary at the prescribed times in a uniformly transversal fashion - cf. Section 5. With this information, the approximate description of Section 5, that was initially used as a guide, is rigorously justified and can be used to quantify (6.1).

6.1. Back-propagating beams

Given a frame element φ_γ, with $\gamma \in \Gamma_h$, we look for a mode, $+$ or $-$, and a tuple of Gaussian beam parameters

$$\mathcal{S}_\gamma^h = (\omega_\gamma^h, a_\gamma^h, \xi_\gamma^h, \mathcal{A}_\gamma^h, \mathcal{M}_\gamma^h),$$

such that (6.1) holds. We write explicitly

$$\varphi_\gamma(t, x_*) = 2^{(j+1/2)d/2} \cdot \exp\Big[i4^j \big(\tilde{\xi}_{j,k} \left(t - 2^{-j}\lambda_1, x_* - 2^{-j}\lambda_*\right) \\ + i\pi \left((t - 2^{-j}\lambda_1)^2 + \left| x_* - 2^{-j}\lambda_* \right|^2 \right) \big) \Big]$$

and compare this expression to (5.4). We want to construct a beam such that:

$$A_\gamma(t_\gamma) = 2^{(j+1/2)d/2}, \tag{6.2}$$

$$\omega_\gamma L_\gamma(t, x_*) = 4^j \tilde{\xi}_{j,k} \left(t - 2^{-j}\lambda_1, x_* - 2^{-j}\lambda_*\right), \tag{6.3}$$

$$\omega_\gamma Q_\gamma(t, x_*) = i\pi 4^j \left((t - 2^{-j}\lambda_1)^2 + \left| x_* - 2^{-j}\lambda_* \right|^2 \right). \tag{6.4}$$

Step 1. *Choice of mode and scale*. Recall that by the non-tangential propagation estimate - cf. (4.7) - $(\xi_{j,k})_1 \neq 0$. Let $\varsigma := \operatorname{sign}((\xi_{j,k})_1) \in \{-1, 1\}$. Note that in the asymptotic expansion in (5.4), the first component of the linear part of the phase is given by (5.1), which is negative for a $+$ beam and positive for a $-$ one. Motivated by this fact, if $\varsigma = -1$ we construct a $+$ mode, while if $\varsigma = 1$ we construct a $-$ mode. Second, we choose the scale parameter as $\omega_\gamma^{\pm,h} = 4^j$. Having made these choices, we ease the notation dropping the superscripts $h, \pm$.
Step 2. *Definition of the boundary intersection time*. We first define the time instant

$$t_\gamma = 2^{-j}\lambda_1. \tag{6.5}$$

The center of the Gaussian beam under construction is to intersect the boundary $\{x_1 = 0\}$ at time t_γ. Note that, due to (4.5), $t_\gamma \in [C_{h,\inf}, C_{h,\sup}]$, and the constants $C_{h,\inf}, C_{h,\sup}$ depend only on the boundary value h, but not on γ.

In the following steps, we define functions $(x(t), p(t), M(t), A(t))$ as solutions of the ODEs in (2.7), (2.9) and (2.10) by specifying adequate initial conditions at time $t = t_\gamma$. Later we define $\mathcal{S}_\gamma^h$ by inspecting $(x(t), p(t), M(t), A(t))$ at time $t = 0$. To this end, we use the description of a beam given in Lemma 5.1. We first aim to match the function $L_\gamma(t, x_*)$ in (5.2) to the linear part of the phase in (2.14).
Step 3. *Definition of* $(x(t), p(t))$. Let $(x, p) : \mathbb{R} \to \mathbb{R}^{2d}$ be the solution of the Hamiltonian flow, cf. (2.7), with initial condition at $t = t_\gamma$ described as follows. For x we simply set:

$$x|_{t=t_\gamma} = (0, 2^{-j}\lambda_*). \tag{6.6}$$

This agrees with our intention that the Gaussian beam under construction Φ_γ intersects the boundary at time t_γ.

With these choices,

$$(t_\gamma, x_*(t_\gamma)) = 2^{-j}\lambda. \tag{6.7}$$

For p we need to specify:

$$p|_{t=t_\gamma} = ((p|_{t=t_\gamma})_1, (p|_{t=t_\gamma})_*).$$

We first define $(p|_{t=t_\gamma})_*$ by

$$(p|_{t=t_\gamma})_* = (\tilde{\xi}_{j,k})_*, \tag{6.8}$$

where $\tilde{\xi}_{j,k}$ is given by (2.1). Second, we define $(p|_{t=t_\gamma})_1$ as

$$(p|_{t=t_\gamma})_1 = (-\varsigma) \cdot \sqrt{\frac{(\tilde{\xi}_{j,k})_1^2}{c(x|_{t=t_\gamma})^2} - \left|(\tilde{\xi}_{j,k})_*\right|^2}. \tag{6.9}$$

Note that $(p|_{t=t_\gamma})_1$ is well-defined because of the grazing ray condition. Indeed, by (4.6),

$$\frac{\left|(\tilde{\xi}_{j,k})_1\right|^2}{c(x(t_\gamma))^2} \geq \left(C_{\text{graz}} + \left|(\tilde{\xi}_{j,k})_*\right|\right)^2 \geq C_{\text{graz}}^2 + \left|(\tilde{\xi}_{j,k})_*\right|^2. \tag{6.10}$$

In addition,

$$\left|p|_{t=t_\gamma}\right|^2 = \left|(p|_{t=t_\gamma})_1\right|^2 + \left|(p|_{t=t_\gamma})_*\right|^2 = \frac{(\tilde{\xi}_{j,k})_1^2}{c(x|_{t=t_\gamma})^2}. \tag{6.11}$$

With these choices, since ς has a sign opposite to the mode of the beam under construction,

$$\tau_\gamma = -H(x|_{t=t_\gamma}, p|_{t=t_\gamma}) = \varsigma \cdot c(x(t_\gamma)) \left|p|_{t=t_\gamma}\right| = \varsigma \cdot \left|(\tilde{\xi}_{j,k})_1\right| = (\tilde{\xi}_{j,k})_1. \tag{6.12}$$

Consequently

$$(\tau_\gamma, (p|_{t=t_\gamma})_*) = \tilde{\xi}_{j,k}, \tag{6.13}$$

and therefore the linear part of the phase of the boundary restriction of the beam under construction - as a function of (t, x_*) and according to the approximate description in Lemma 5.1 - coincides with the linear part of the phase of $\varphi_\gamma(t, x_*)$. Moreover, we note the following.

Claim 6.1. *The flow $(x(t), p(t)) = (x_\gamma(t), p_\gamma(t))$ defined in Step 3 satisfies:*

$$\dot{x}_{\gamma,1}(t_\gamma) > 0, \text{ and } \dot{x}_{\gamma,1}(t_\gamma) \gtrsim 1, \tag{6.14}$$

$$\left|p_{\gamma,1}(t_\gamma)\right| \gtrsim \left|p_\gamma(t_\gamma)\right|, \tag{6.15}$$

where the implied constants are uniform for $\gamma \in \Gamma_h$.

Proof. From (6.11) we see that $\left|p_\gamma(t_\gamma)\right| \lesssim 1$. In addition, (6.9) and (6.10) imply that $\left|p_{\gamma,1}(t_\gamma)\right| \gtrsim 1$, so the claim in (6.15) follows. For (6.14), we use one of Hamilton's equations

$$\dot{x}_{\gamma,1}(t_\gamma) = \partial_{p_1} H(t_\gamma) = (-\varsigma) \cdot c\left(x_\gamma(t_\gamma)\right) \frac{p_{\gamma,1}(t_\gamma)}{|p_\gamma(t_\gamma)|}.$$

Inspecting the sign of (6.9) we see that $\dot{x}_{\gamma,1}(t_\gamma) > 0$. In addition, by the assumptions on the velocity (1.10) and (6.15),

$$\left|\dot{x}_{\gamma,1}(t_\gamma)\right| \geq C_{\mathrm{vel}} \frac{\left|p_{\gamma,1}(t_\gamma)\right|}{|p_\gamma(t_\gamma)|} \gtrsim 1. \quad \square$$

Step 4. *Definition of $M(t_\gamma)$*. Let

$$\widetilde{M}_\gamma = 2\pi i I_d,$$

and let $M(t_\gamma) \in \mathbb{C}^{d\times d}$ be the unique symmetric matrix that solves the following system of equations:

$$\begin{cases} \widetilde{M}_{\gamma,11} &= \dot{x}_\gamma(t_\gamma) \cdot M(t_\gamma)\dot{x}_\gamma(t_\gamma) - \dot{p}_\gamma(t_\gamma) \cdot \dot{x}_\gamma(t_\gamma), \\ \widetilde{M}_{\gamma,1k} &= \dot{p}_{\gamma,k}(t_\gamma) - \sum_{n=1}^d \left(M(t_\gamma)\right)_{kn} \dot{x}_{\gamma,n}(t_\gamma), \qquad k = 2, \dots d, \\ \widetilde{M}_{\gamma,kl} &= \left(M(t_\gamma)\right)_{kl}, \qquad k, l = 2, \dots d. \end{cases} \tag{6.16}$$

We now check that $M(t_\gamma)$ is indeed well-defined.

Claim 6.2. *The system* (6.16) *has a unique symmetric solution $M(t_\gamma)$. Moreover, there exist constants $C_1, C_2 > 0$ - independent of γ - such that $\|M(t_\gamma)\| \leq C_1$ and $\Im M(t_\gamma) \geq C_2 \cdot I_d$.*

We postpone the proof of the claim to Section 8.1, so as not to interrupt the flow of the construction.

Step 5. *Definition of $M(t)$*. We let $M(t)$ be the solution of (2.9) with initial condition at time $t = t_\gamma$ given by the matrix $M(t_\gamma)$ from Step 4. Due to Claim 6.2, this is a valid initial condition - cf. Section 2.4.

Step 6. *Definition of $A(t)$*. Let $A(t)$ be the solution to (2.10) with initial condition:

$$A(t_\gamma) = 2^{(j+1/2)\frac{d}{2}}. \tag{6.17}$$

Step 7. *Definition of $\mathcal{S}_\gamma^{h,+}$ and $\mathcal{S}_\gamma^{h,-}$*. We recall the decomposition $\Gamma_h = \Gamma_h^+ \cup \Gamma_h^-$ in (4.8) and define two sets of GB parameters

$$\left\{\mathcal{S}_\gamma^{h,+} : \gamma \in \Gamma_h^+\right\}, \qquad \left\{\mathcal{S}_\gamma^{h,-} : \gamma \in \Gamma_h^-\right\},$$

with $\mathcal{S}_\gamma^{h,\pm} = (\omega_\gamma^{\pm,h}, a_\gamma^{\pm,h}, \xi_\gamma^{\pm,h}, \mathcal{A}_\gamma^{\pm,h}, \mathcal{M}_\gamma^{\pm,h})$ in the following way:

$$\begin{cases} \omega_\gamma^{\pm,h} = 4^j, & a_\gamma^{\pm,h} = x^\pm(0), \qquad \xi_\gamma^{\pm,h} = \dfrac{1}{2\pi} 4^j p^\pm(0), \\ \mathcal{A}_\gamma^{\pm,h} = 4^{-j\frac{d}{4}} A^\pm(0), & \qquad \mathcal{M}_\gamma^{\pm,h} = \dfrac{1}{2\pi} M^\pm(0). \end{cases} \tag{6.18}$$

The values for $\mathcal{S}_\gamma^{h,\pm}$ are chosen so that if we define the functions $(x^\pm(t), p^\pm(t), M^\pm(t), A^\pm(t))$ by imposing initial conditions at time $t = 0$ as described in Section 2.4, they will satisfy (6.6), (6.7), (6.13) and (6.17) at time $t = t_\gamma$. As a result of the construction, (6.2), (6.3), (6.4) are satisfied.

Remark 6.3. The function $\mathcal{S}_\gamma^{h,+}$ is defined only on Γ_h^+. According to the conventions in Section 2.4, it would be possible to associate with such map both a family of $+$ beams and a family of $-$ beams. However, we are only interested in the corresponding family of $+$ beams, because, as we show below, these satisfy the approximation property in (6.1). A similar remark applies to $\mathcal{S}_\gamma^{h,-}$.

Remark 6.4. The choice of sign in (6.9) is instrumental to construct a parametrix for the Dirichlet problem on the right-half space (see Theorem 6.7 below). For the left-half space, the opposite sign should be used in (6.9), leading to a different sign in Claim 6.1.

6.2. *Analysis of the back-propagated parameters*

We now analyze the properties of the previous construction. We first state the following fundamental property.

Theorem 6.5 *(Well-spreadness). Each of the two families of back-propagated parameters constructed in Section 6.1,* $\Upsilon^{h,+} = \left\{\mathcal{S}_\gamma^{h,+} : \gamma \in \Gamma_h^+\right\}$, $\Upsilon^{h,-} = \left\{\mathcal{S}_\gamma^{h,-} : \gamma \in \Gamma_h^-\right\}$ *is a well-spread set of Gaussian beam parameters.*

The proof of Theorem 6.5 is quite technical and we postpone it to Section 8. We now analyze the fine properties of the matching procedure.

Theorem 6.6 *(Transversal boundary intersection). Each family of beams* $\{\Phi_\gamma^{h,\pm} : \gamma \in \Gamma_h^\pm\}$ *intersects the boundary* $\{x_1 = 0\}$ *at times* $\{t_\gamma : \gamma \in \Gamma_h^\pm\}$ *- given by* (6.5) *- in a uniformly transversal fashion.*

Proof. By (6.6), $x_{\gamma,1}^h(t_\gamma) = 0$. The uniform transversality at the boundary intersection is proved in Claim 6.1 - see (6.15). □

Theorem 6.7 *(Rightwards propagation). The spatial centers of the beams* $\{\Phi_\gamma^{h,\pm} : \gamma \in \Gamma_h^\pm\}$ *are uniformly away from the right-half plane* $\mathbb{R}_+^d$ *at time* $t = 0$*. More precisely, there exists a constant* $\epsilon > 0$ *such that, for all* $\gamma \in \Gamma_h^\pm$,

$$x_{\gamma,1}^{h,\pm}(0) \le -\epsilon. \tag{6.19}$$

Proof. By Theorem 6.6, $x_{\gamma,1}(t_\gamma) = 0$ and $\left|p_{\gamma,1}(t_\gamma)\right| \gtrsim \left|p_\gamma(t_\gamma)\right|$. In addition, $t_\gamma = 2^{-j}\lambda_1 \in [C_{h,\inf}, C_{h,\sup}] \subseteq [0, T]$ by the *approximate compatibility condition* (4.5). Let us write

$$x_{\gamma}^{h,\pm}(t) = x_{\gamma}^{h,\pm}(t_{\gamma}) + (t - t_{\gamma})b(t),$$

with b smooth. The *cone condition* (4.3) implies that $|b_1(t)| \gtrsim 1$, for $t \in [0, T]$. In addition, $b_1(t_\gamma) = \dot{x}_{\gamma,1}^{h,\pm}(t_\gamma) > 0$ by Claim 6.1. Hence, $b_1 > 0$ on $[0, T]$ and, moreover, $b_1(t) \gtrsim 1$ for all $t \in [0, T]$. Second, the *approximate compatibility condition* (4.5) implies that $t_\gamma = 2^{-j}\lambda_1 \gtrsim 1$. Therefore,

$$-x_{\gamma,1}^{h,\pm}(0) = t_{\gamma} \cdot b_1(0) \gtrsim 1,$$

as claimed. □

Theorem 6.8 *(Beams match frame elements on the boundary). When restricted to the boundary $\{x_1 = 0\}$, the beams $\{\Phi_\gamma^{h,\pm} : \gamma \in \Gamma_h^{\pm}\}$ match the frame elements in the following sense. Let $\eta^1 \in C^\infty(\mathbb{R})$ be compactly supported. Let $\eta^2 \in C^\infty(\mathbb{R}^{d-1})$ be a smooth function supported on $B_2(0)$ that is $\equiv 1$ on $B_1(0)$, and let $\eta_\gamma^2(x_*) = \eta(x_* - x_{\gamma,*}(t_\gamma))$.*

• Local description*:*

$$\left(\Phi_{\gamma}^{h,\pm}(t, 0, x_*) - \varphi_{\gamma}(t, x_*)\right) \cdot \eta^1(t) \cdot \eta_{\gamma}^2(x_*) = \left(4^j \cdot R_{\gamma}^1(t, x_*) + R_{\gamma}^2(t, x_*)\right) \cdot \varphi_{\gamma}(t, x_*),$$

with $\gamma = (j, k, \lambda) \in \Gamma_h^{\pm}$, $R^1 = \mathbb{O}_{\geq}^3(\{0\}, \Upsilon^{h,\pm})$ and $R^2 = \mathbb{O}_{\geq}^1(\{0\}, \Upsilon^{h,\pm})$.

• Global description*:*

$$\left(\Phi_{\gamma}^{h,\pm}(t, 0, x_*) - \varphi_{\gamma}(t, x_*)\right) \cdot \eta^1(t) \cdot (1 - \eta_{\gamma}^2(x_*)) = \tilde{\Phi}_{\gamma}^{\pm}(t, 0, x_*) \cdot R_{\gamma}^3(t, x_*),$$

with $\tilde{\Upsilon}^{\pm} \equiv \{\tilde{\mathcal{S}}_\gamma : \gamma \in \Gamma_h^{\pm}\}$ well-spread sets of GB parameters, $\tilde{\Phi}_\gamma^{\pm}$ the corresponding beams and $R^3 = \mathbb{O}_{\geq}^1(\{0\}, \tilde{\Upsilon}^{\pm})$.

Remark 6.9. We stress that here the time variable t is not considered as an evolution variable; rather (t, x_*) functions as a spatial variable. In accordance, $\{0\}$ is the time-evolution set in the $\mathbb{O}_{\geq}$ notation.

Proof of Theorem 6.8. We invoke Lemmas 5.1 and 5.2. The corresponding hypothesis is satisfied, thanks to Theorems 6.5 and 6.6. We use the notation $\mathcal{S}_\gamma^{h,\pm} = (\omega_\gamma^{\pm,h}, a_\gamma^{\pm,h}, \xi_\gamma^{\pm,h}, \mathcal{A}_\gamma^{\pm,h}, \mathcal{M}_\gamma^{\pm,h})$.

For the *local description*, due to Theorem 6.5, we can invoke Lemma 5.1. We substitute the values of the beam parameters defined in Section 6.1 into (5.4) - cf. (6.2), (6.3), (6.4) and obtain:

$$\Phi_{\gamma}^{h,\pm}(t, 0, x_*) = \varphi_{\gamma}(t, x_*) \left(1 + R_{\gamma}(t)\right) e^{i \cdot 4^j \cdot \Theta_{\gamma}(t, x_*)}, \qquad (t, x_*) \in \mathbb{R}_T^d,$$

with R_γ and Θ_γ as in Lemma 5.1. Second, we note that

$$e^{i \cdot 4^j \cdot \Theta_{\gamma}(t, x_*)} \cdot \eta^1(t) \cdot \eta_{\gamma}^2(x_*) = 1 + 4^j \cdot R_{\gamma}'(t, x_*),$$

with $R' = \mathbb{O}_{\geq}^3(\{0\}, \Upsilon^{h,\pm})$, and the conclusion follows.

For the *global description*, with the notation of Lemma 5.2,

$$\Phi_\gamma^{h,\pm}(t,0,x_*) = A_\gamma(t_\gamma)\exp\left[i4^j(L_\gamma(t,x_*) + i\ell\left((t-t_\gamma)^2 + \left|x_* - x_{\gamma,*}(t_\gamma)\right|^2\right))\right]\cdot R_\gamma(t,x_*).$$

Substituting the values of the parameters defined in Section 6 - cf. (6.7) and (6.13) - we obtain

$$\Phi_\gamma^{h,\pm}(t,0,x_*) = 2^{(j+1/2)\frac{d}{2}} e^{2\pi i\left((t,x_*)-2^{-j}\lambda\right)\xi_{j,k} - \ell 4^j\left|(t,x_*)-2^{-j}\lambda\right|^2}\cdot R_\gamma(t,x_*).$$

We let $R_\gamma^3(t,x_*) := \eta^1(t)(1-\eta_\gamma^2(x_*))R_\gamma(t,x_*)$ and

$$\tilde{\mathcal{S}}_\gamma^{\pm} = (4^j, 2^{-j}\lambda, \xi_{j,k}, 2^{\frac{d}{4}}, i\tfrac{\ell}{\pi}I_d), \qquad \gamma = (j,k,\lambda)\in\Gamma_h^{\pm}.$$

It is straightforward to verify that this defines a well-spread set of GB parameters. Indeed, for $\gamma\in\Gamma_h^{\pm}$, the tuple $\tilde{\mathcal{S}}_\gamma^{\pm}$ is very similar to the standard one $\mathcal{S}_\gamma^{st}$, defined in (2.15): the only difference is that, in the new set, the standard matrix element $\mathcal{M}_\gamma = iI_d$ is replaced by $i\frac{\ell}{\pi}I_d$, with $\ell > 0$ a constant. □

7. Parametrix estimates for the Dirichlet problem

Finally, we derive the parametrix for the boundary Dirichlet problem and give suitable estimates.

Theorem 7.1. *With the assumptions and notation from Section 4, let $u : [0,T]\times\mathbb{R}_+^d \to \mathbb{C}$ be the (weak) solution to the problem:*

$$\begin{cases} \partial_t^2 u(t,x) - c(x)^2\Delta_x u(t,x) = 0, & t\in[0,T], x\in\mathbb{R}_+^d,\\ u(0,x) = u_t(0,x) = 0, & x\in\mathbb{R}_+^d,\\ u(t,0,y) = h_{cut}(t,y), & t\in[0,T], y\in\mathbb{R}^{d-1}. \end{cases} \tag{7.1}$$

Let $\tilde{h} = \sum_{\gamma\in\Gamma_h}\tilde{h}_\gamma\varphi_\gamma$ be the truncated frame expansion of h defined in Section 4.2 and consider the GB parameters $\mathcal{S}_\gamma^{h,\pm}$ constructed in Section 6. Let $\tilde{u}$ be defined as

$$\tilde{u} = \sum_{\gamma\in\Gamma_h^+}\tilde{h}_\gamma\Phi_\gamma^{h,+} + \sum_{\gamma\in\Gamma_h^-}\tilde{h}_\gamma\Phi_\gamma^{h,-}. \tag{7.2}$$

Then

$$\|\tilde{u} - u\|_{C^0([0,T],H^1(\mathbb{R}_+^d))\cap C^1([0,T],L^2(\mathbb{R}_+^d))} \le C_T\|h\|_{H^{1/2}(\mathbb{R}^d)}.$$

In particular, in the highly oscillatory regime: $\hat{h}(\xi) = 0$ for $|\xi|\le\xi_{\min}$, we obtain

$$\|\tilde{u} - u\|_{C^0([0,T],H^1(\mathbb{R}_+^d))\cap C^1([0,T],L^2(\mathbb{R}_+^d))} \le C_T\cdot\xi_{\min}^{-1/2}\cdot\|h\|_{H^1(\mathbb{R}^d)}.$$

(Here, $\hat{h}$ denotes the Fourier transform of h in the full (t,y) variable.)

Remark 7.2. The problem in (7.1) is well-posed because h_{cut} satisfies the compatibility condition $h_{cut}(0,\cdot)\equiv 0$, cf. Section 4.

To prove Theorem 7.1, we show that the proposed GB solution approximately solves the boundary-value problem and then conclude, by means of energy estimates, that it must suitably approximate the true solution. The results from Section 3, together with the analysis of the back-propagated parameters in Section 6, imply that the wave operator approximately annihilates the GB solution. In the next section, we show that the other conditions of the boundary-value problem are also approximately satisfied.

7.1. Preliminary steps

As a first step towards the proof of Theorem 7.1, we show that the approximate solution in (7.2) satisfies zero boundary conditions, up to the error of the parametrix. More precisely, we have the following lemma.

Lemma 7.3 *(Asymptotic vanishing of the initial conditions). Under the hypothesis of Theorem 7.1, consider the approximate solution defined in* (7.2)*. Then*

$$\|\tilde{u}|_{t=0}\|_{H^1(\mathbb{R}^d_+)},\ \|\partial_t\tilde{u}|_{t=0}\|_{L^2(\mathbb{R}^d_+)}\lesssim C_T\|h\|_{H^{1/2}(\mathbb{R}^d)}.$$

Proof. We use the short notation $u=\sum_{\gamma\in\Gamma_h}\tilde{h}_\gamma\Phi^h_\gamma$, with the understanding that Φ^h_γ is a $+$ mode for $\gamma\in\Gamma_h^+$ and a $-$ mode for $\gamma\in\Gamma_h^-$. We drop the $\pm$ superscripts on solutions to the defining ODEs, with a similar convention. At time $t=0$ we have

$$\tilde{u}|_{t=0}=\sum_{\gamma\in\Gamma_h}\tilde{h}_\gamma\Phi^h_\gamma(0,\cdot), \tag{7.3}$$

$$\partial_t\tilde{u}|_{t=0}=\sum_{\gamma\in\Gamma_h}\tilde{h}_\gamma\partial_t\Phi^h_\gamma(0,\cdot). \tag{7.4}$$

By Theorem 6.7, the centers of beams $\Phi^{h,\pm}_\gamma$ are away from the boundary $\{x_1=0\}$ at initial time; we let $\epsilon>0$ be such that (6.19) holds.

Step 1. *Localization*. Intuitively, (6.19) means that, at time $t=0$, the right-half space is away from the wave-front set of the solution, and the parametrix is micro-locally of lower order. To formalize this reasoning, let us consider a smooth cut-off function $\eta:\mathbb{R}^d\to[0,1]$, such that

$$\eta(x)=\begin{cases}0, & x_1\le-\epsilon,\\ 1, & x_1\ge-\epsilon/2.\end{cases}$$

We also define $\eta_\gamma:=\eta$ for all $\gamma\in\Gamma_h$. By (6.19), η vanishes near $x^h_\gamma(0)$ and, therefore,

$$\{\eta_\gamma:\gamma\in\Gamma_h^\pm\}=\mathbb{O}^k_\geq\left(\Upsilon^{h,\pm},\{0\}\right),\qquad \text{for all } k\ge 0. \tag{7.5}$$

Step 2. *The H^1 norm of* (7.3). We use the fact that the back-propagated parameters are well-spread - Theorem 6.5, the Bessel bounds - Theorem 3.5, and (7.5) with $k=1$ to estimate

$$\Big\| \sum_{\gamma\in\Gamma_h} \tilde{h}_\gamma \Phi^h_\gamma(0,\cdot)\Big\|^2_{H^1(\mathbb{R}^d_+)} \le \Big\| \sum_{\gamma\in\Gamma_h} \tilde{h}_\gamma \Phi^h_\gamma(0,\cdot)\eta\Big\|^2_{H^1(\mathbb{R}^d)}$$
$$\le C_T \sum_{\gamma\in\Gamma_h} 4^{j(2-1)}|\tilde{h}_\gamma|^2 \le C_T\|h\|^2_{H^{1/2}}.$$

Step 3. *The L^2 norm of* (7.4). Using (3.3) we see that

$$\partial_t\Phi^h_\gamma(0,x) = \left(F^1_\gamma(x) + 4^j F^2_\gamma(x)\right)\Phi^{h,\pm}_\gamma,$$

with $F^1, F^2 = \mathbb{O}^0_{\ge}(\Upsilon^{h,\pm},\{0\})$. Combining this with (7.5), we can proceed as in Step 2 to deduce that

$$\| \sum_{\gamma\in\Gamma_h} \tilde{h}_\gamma \partial_t\Phi^h_\gamma(0,\cdot)\|_{L^2(\mathbb{R}^d_+)} \le \| \sum_{\gamma\in\Gamma_h} \tilde{h}_\gamma \partial_t\Phi^h_\gamma(0,\cdot)\eta\|_{L^2(\mathbb{R}^d)} \le C_T\|h\|_{H^{1/2}}.$$

This completes the proof. □

Theorem 7.4 *(Boundary conditions are asymptotically satisfied). Under the hypothesis of Theorem 7.1, the Gaussian beam solution $\tilde{u}$ satisfies:*

$$\|\tilde{u}(\cdot,0,\cdot) - h_{cut}\|_{H^1([0,T]\times\mathbb{R}^{d-1})} \le C_T\|h\|_{H^{1/2}(\mathbb{R}^d)}.$$

Proof. We use the same short-hand notation as in the proof of Lemma 7.3. According to the definitions,

$$\tilde{h}(t,x_*) = \sum_{\gamma\in\Gamma_h} \tilde{h}_\gamma\varphi_\gamma(t,x_*),$$
$$\tilde{u}(t,0,x_*) = \sum_{\gamma\in\Gamma_h} \tilde{h}_\gamma\Phi^h_\gamma(t,0,x_*), \qquad t\in\mathbb{R}, x_*\in\mathbb{R}^{d-1}.$$

Therefore,

$$\|\tilde{u}(\cdot,0,\cdot) - h_{cut}\|_{H^1([0,T]\times\mathbb{R}^{d-1})}$$
$$\le \|h_{cut} - \tilde{h}\|_{H^1([0,T]\times\mathbb{R}^{d-1})} + \Big\|\sum_{\gamma\in\Gamma_h}\tilde{h}_\gamma\left(\varphi_\gamma - \Phi^h_\gamma(\cdot,0,\cdot)\right)\Big\|_{H^1([0,T]\times\mathbb{R}^{d-1})}.$$

By (4.4), the first term in the last equation is suitably bounded. Let us focus on the second term.

We invoke Theorem 6.8. Let $\eta^1 \in C^\infty(\mathbb{R})$ be a smooth compactly-supported cut-off window such that $\eta^1 \equiv 1$ on $[0,T]$ and $\eta^2 \in C^\infty(\mathbb{R}^{d-1})$ a smooth function supported on $B_2(0)$ that is $\equiv 1$ on $B_1(0)$. We write $\eta^2_\gamma(x_*) = \eta^2(x_* - x_{\gamma,*}(t_\gamma))$. With the notation of Theorem 6.8,

$$\begin{aligned}
\Big\| \sum_{\gamma\in\Gamma_h} \tilde{h}_\gamma \big(\varphi_\gamma - \Phi^h_\gamma(\cdot,0,\cdot)\big) \Big\|_{H^1([0,T]\times\mathbb{R}^{d-1})} &\le \Big\| \sum_{\gamma\in\Gamma_h} \tilde{h}_\gamma \big(\varphi_\gamma - \Phi^h_\gamma(\cdot,0,\cdot)\big)\eta^1 \Big\|_{H^1(\mathbb{R}^d)} \\
&\le \Big\| \sum_{\gamma\in\Gamma_h} \tilde{h}_\gamma \big(\varphi_\gamma - \Phi^h_\gamma(\cdot,0,\cdot)\big)\eta^1 \left(1-\eta^2_\gamma\right) \Big\|_{H^1(\mathbb{R}^d)} \\
&\quad + \Big\| \sum_{\gamma\in\Gamma_h} \tilde{h}_\gamma \big(\varphi_\gamma - \Phi^h_\gamma(\cdot,0,\cdot)\big)\eta^1 \eta^2_\gamma \Big\|_{H^1(\mathbb{R}^d)} \\
&\le \Big\| \sum_{\gamma\in\Gamma_h} \tilde{h}_\gamma \varphi_\gamma \left(4^j R^1_\gamma + R^2_\gamma\right) \Big\|_{H^1(\mathbb{R}^d)} + \Big\| \sum_{\gamma\in\Gamma_h} \tilde{h}_\gamma \tilde{\Phi}_\gamma(\cdot,0,\cdot) R^3_\gamma \Big\|_{H^1(\mathbb{R}^d)}.
\end{aligned}$$

We use the information on the vanishing orders of R^k, $k = 1, 2, 3$, the fact that the beams $\{\tilde{\Phi}_\gamma\}$ are well-spread, and the Bessel bounds from Theorem 3.5 - with (t, x_*) as integration variable instead of x - to conclude see that the remaining terms are dominated by $\|h\|_{H^{1/2}(\mathbb{R}^d)}$. This completes the proof. □

7.2. *Proof of the main result*

Proof of Theorem 7.1. The function $v := \tilde{u} - u$ solves the problem:

$$\begin{cases}
\partial_t^2 v(t,x) - c(x)^2 \Delta_x v(t,x) = f(t,x), & t\in[0,T],\, x\in\mathbb{R}^d_+, \\
v(0,x) = \tilde{u}(0,x), & x\in\mathbb{R}^d_+, \\
v_t(0,x) = \tilde{u}_t(0,x), & x\in\mathbb{R}^d_+, \\
v(t,0,x_*) = \tilde{u}(t,0,x_*) - h_{cut}(t,x_*), & t\in[0,T],\, x_*\in\mathbb{R}^{d-1},
\end{cases}$$

where $f(t,x) = \partial_t^2 \tilde{u}(t,x) - c(x)^2 \Delta_x \tilde{u}(t,x)$. By the energy estimates for the wave equation [27, 26,45],

$$\begin{aligned}
&\sup_{t\in[0,T]} \|v(t,\cdot)\|_{H^1(\mathbb{R}^d_+)} + \sup_{t\in[0,T]} \|\partial_t v(t,\cdot)\|_{L^2(\mathbb{R}^d_+)} \\
&\quad \le C_T \Big(\|\tilde{u}(0,\cdot)\|_{H^1(\mathbb{R}^d_+)} + \|\tilde{u}_t(0,\cdot)\|_{L^2(\mathbb{R}^d_+)} + \sup_{t\in[0,T]} \|f(t,\cdot)\|_{L^2(\mathbb{R}^d_+)} \\
&\qquad + \|\tilde{u}(\cdot,0,\cdot) - h_{cut}\|_{H^1([0,T]\times\mathbb{R}^{d-1})} \Big).
\end{aligned}$$

The term involving f can be estimated by Theorems 3.8 and 6.5 as

$$\sup_{t\in[0,T]} \|f(t,\cdot)\|^2_{L^2} \le C_T \sum_{\gamma\in\Gamma_h} 4^j |\tilde{h}_\gamma|^2 \le C_T \|h\|^2_{H^{1/2}(\mathbb{R}^d)},$$

while the other three terms are similarly bounded, by Lemma 7.3 and Theorem 7.4. This completes the proof. □

8. Proofs related to the back-propagated parameters

This section is devoted to pending proofs related to Section 6.

8.1. Proof of Claim 6.2

Proof. We use the notation of Section 6.1.

Step 1. *Existence and uniqueness*. In compact notation, we look for a symmetric matrix $M(t_\gamma) \in \mathbb{C}^{d\times d}$ such that

$$\begin{bmatrix} \widetilde{M}_{\gamma,11} & \widetilde{M}^t_{\gamma,1*} \\ \widetilde{M}_{\gamma,1*} & \left(M_\gamma(t_\gamma)\right)_{**} \end{bmatrix} = \begin{bmatrix} 2\pi i & 0 \\ 0 & 2\pi i I_{d-1} \end{bmatrix} \tag{8.1}$$

where

$$\widetilde{M}_{\gamma,11} = \dot{x}_\gamma(t_\gamma)\cdot M(t_\gamma)\dot{x}_\gamma(t_\gamma) - \dot{p}_\gamma(t_\gamma)\cdot\dot{x}_\gamma(t_\gamma), \tag{8.2}$$

$$\widetilde{M}_{\gamma,1*} = \left(\dot{p}_\gamma(t_\gamma) - M(t_\gamma)\dot{x}_\gamma(t_\gamma)\right)_*. \tag{8.3}$$

We first assume that we have such a matrix $M(t_\gamma)$ and deduce the values of its entries. From (8.1) we see that

$$(M(t_\gamma))_{**} = 2\pi i I_{d-1}. \tag{8.4}$$

Using this together with (8.3) and (8.1) we see that

$$\left(M(t_\gamma)\right)_{k1}\dot{x}_{\gamma,1}(t_\gamma) = \dot{p}_{\gamma,k}(t_\gamma) - 2\pi \mathrm{i}\dot{x}_{\gamma,k}(t_\gamma), \qquad k = 2,\dots d.$$

Since, by (6.14), $\dot{x}_{\gamma,1}(t_\gamma) \neq 0$ and $M(t_\gamma)$ is symmetric, we can solve

$$\left(M(t_\gamma)\right)_{1*} = \left(\dot{x}_{\gamma,1}(t_\gamma)\right)^{-1}\left(\dot{p}_{\gamma,*}(t_\gamma) - 2\pi\mathrm{i}\dot{x}_{\gamma,*}(t_\gamma)\right). \tag{8.5}$$

We now compare the (1, 1) entries in (8.1) and use (8.4) and (8.5) together with (8.2) to obtain

$$\begin{aligned} 2\pi i &= \dot{x}_\gamma(t_\gamma)\cdot M(t_\gamma)\dot{x}_\gamma(t_\gamma) - \dot{p}_\gamma(t_\gamma)\cdot\dot{x}_\gamma(t_\gamma) \\ &= \left|\dot{x}_{\gamma,1}(t_\gamma)\right|^2\left(M(t_\gamma)\right)_{11} + 2\dot{x}_{\gamma,1}(t_\gamma)\cdot\left(M(t_\gamma)\right)_{1*}\dot{x}_{\gamma,*}(t_\gamma) \\ &\quad + \dot{x}_{\gamma,*}(t_\gamma)\cdot\left(M(t_\gamma)\right)_{**}\dot{x}_{\gamma,*}(t_\gamma) - \dot{p}_\gamma(t_\gamma)\cdot\dot{x}_\gamma(t_\gamma) \\ &= \left|\dot{x}_{\gamma,1}(t_\gamma)\right|^2\left(M(t_\gamma)\right)_{11} + 2\left(\dot{p}_{\gamma,*}(t_\gamma) - 2\pi i\dot{x}_{\gamma,*}(t_\gamma)\right)\cdot\dot{x}_{\gamma,*}(t_\gamma) \\ &\quad + 2\pi i\left|\dot{x}_{\gamma,*}(t_\gamma)\right|^2 - \dot{p}_\gamma(t_\gamma)\cdot\dot{x}_\gamma(t_\gamma) \\ &= \left|\dot{x}_{\gamma,1}(t_\gamma)\right|^2\left(M(t_\gamma)\right)_{11} + 2\dot{p}_{\gamma,*}(t_\gamma)\cdot\dot{x}_{\gamma,*}(t_\gamma) \\ &\quad - 2\pi i\left|\dot{x}_{\gamma,*}(t_\gamma)\right|^2 - \dot{p}_\gamma(t_\gamma)\cdot\dot{x}_\gamma(t_\gamma). \end{aligned}$$

Using again that, by (6.14), $\dot{x}_{\gamma,*}(t_\gamma) \neq 0$ and we conclude that

$$\left(M(t_\gamma)\right)_{11} = \left|\dot{x}_{\gamma,1}(t_\gamma)\right|^{-2}\left(2\pi i\left(1 + \left|\dot{x}_{\gamma,*}(t_\gamma)\right|^2\right) - 2\dot{p}_{\gamma,*}(t_\gamma)\cdot\dot{x}_{\gamma,*}(t_\gamma) + \dot{p}_\gamma(t_\gamma)\cdot\dot{x}_\gamma(t_\gamma)\right). \tag{8.6}$$

Hence, the matrix $M(t_\gamma)$ is completely determined by the desired conditions. Let us define $M(t_\gamma)$ by (8.4), (8.5) and (8.6) and the requirement of symmetry. We see that such matrix solves (6.16).

Step 2. *Positivity and bounds*. Inspecting (8.4), (8.5) and (8.6) and using Claim 6.1 we see that $\|M(t_\gamma)\|$ is bounded uniformly for $\gamma \in \Gamma_h$. According to the definitions, the imaginary part of $M(t_\gamma)$ is of the form

$$\frac{1}{2\pi}\Im M(t_\gamma) = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1d} \\ a_{12} & 1 & 0 & \dots & 0 \\ a_{13} & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & 0 \\ a_{1d} & 0 & 0 & \dots & 1 \end{bmatrix},$$

where

$$a_{11} = |\dot{x}_{\gamma,1}(t_\gamma)|^{-2}(1 + \left|x_*(t_\gamma)\right|^2),$$
$$a_{1k} = -\dot{x}_{\gamma,1}(t_\gamma))^{-1}\dot{x}_{\gamma,k}(t_\gamma), \qquad k = 2, \dots d.$$

Note that

$$a_{11} - a_{12}^2 - \dots - a_{1d}^2 = |\dot{x}_{\gamma,1}(t_\gamma)|^{-2} \gtrsim 1,$$

by (6.14). A linear algebra argument now shows that $\Im(M(t_\gamma))$ is a positive matrix and $\Im(M(t_\gamma)) \gtrsim I_d$, as desired - see [5, Lemma A.1]. □

8.2. *Proof of Theorem 6.5*

The goal of this section is to show that both families of back-propagated GB parameters constructed in Section 6 are well-spread. This involves comparing the constructed maps

$$\mathcal{S}_\gamma^{h,\pm} = (\omega_\gamma^{\pm,h}, a_\gamma^{\pm,h}, \xi_\gamma^{\pm,h}, \mathcal{A}_\gamma^{\pm,h}, \mathcal{M}_\gamma^{\pm,h}), \qquad \gamma \in \Gamma_h^\pm,$$

to the standard one

$$\mathcal{S}_\gamma^{st} = (4^j, 2^{-j}\lambda, \xi_{j,k}, 2^{\frac{d}{4}}, iI_d), \qquad \gamma = (j, k, \lambda) \in \Gamma.$$

We follow the notation of Section 6: when convenient, we drop the superscripts for the functions $x_\gamma^{h,\pm}(t)$, $p_\gamma^{h,\pm}(t), \dots$, writing instead $x_\gamma(t)$, $p_\gamma(t), \dots$ We keep however the superscripts in the tuple of parameters $\mathcal{S}_\gamma^{h,\pm}$ to avoid confusion with the standard one.

Recall from Section 6 that for $\gamma \in \Gamma_h$, $t_\gamma = 2^{-j}\lambda_1$ and that $x_\gamma(t_\gamma) = (0, 2^j\lambda_*)$.

As a preparation for the proof of Theorem 6.5, we show the following.

Lemma 8.1. *For $\gamma = (j, k, \lambda)$, $\gamma' = (j', k', \lambda') \in \Gamma_h^+$:*

$$\left|\xi_{j,k} - \xi_{j',k'}\right|^2 \lesssim \left|4^j p_\gamma^{h,+}(t_\gamma) - 4^{j'} p_{\gamma'}^{h,+}(t_{\gamma'})\right|^2 + 4^{j+j'}|2^{-j}\lambda_* - 2^{-j'}\lambda'_*|^2.$$

An analogous statement holds for Γ_h^-.

Proof. We treat the family Γ_h^+. To further simplify the notation, throughout this proof we write $p_\gamma = p_\gamma(t_\gamma)$, $p_{\gamma,1} = p_{\gamma,1}(t_\gamma)$, $p_{\gamma,*} = p_{\gamma,*}(t_\gamma)$, and $c_\gamma = c(0, 2^{-j}\lambda_*)$. Recall also that $\tau_\gamma =$

$(\tilde{\xi}_{j,k})_1 < 0$ - cf. (6.12) and (6.13). Hence, by (6.8) and (6.9),

$$p_{\gamma,1} = \sqrt{\frac{\tau_\gamma^2}{c_\gamma^2} - p_{\gamma,*}^2} = \frac{|\tau_\gamma|}{c_\gamma}\sqrt{1 - \frac{c_\gamma^2}{\tau_\gamma^2} p_{\gamma,*}^2}, \tag{8.7}$$

and, by (2.1), $(\xi_{j,k})_1 = \frac{4^j}{2\pi}\tau_\gamma$, $(\xi_{j,k})_* = \frac{4^j}{2\pi} p_{\gamma,*}$. With this notation, the estimate we want to prove is:

$$\left|4^j \tau_\gamma - 4^{j'} \tau_{\gamma'}\right|^2 + |4^j p_{\gamma,*} - 4^{j'} p_{\gamma',*}|^2 \lesssim |4^j p_\gamma - 4^{j'} p_{\gamma'}|^2 + 4^{j+j'} |2^{-j}\lambda_* - 2^{-j'}\lambda'_*|^2.$$

Clearly, it suffices to show that

$$\left|4^j \tau_\gamma - 4^{j'} \tau_{\gamma'}\right|^2 \lesssim |4^j p_\gamma - 4^{j'} p_{\gamma'}|^2 + 4^{j+j'} |2^{-j}\lambda_* - 2^{-j'}\lambda'_*|^2. \tag{8.8}$$

Step 1. *We show that* $\left|4^j \tau_\gamma - 4^{j'} \tau_{\gamma'}\right|^2 \lesssim \left|4^j \frac{\tau_\gamma}{c_\gamma} - 4^{j'} \frac{\tau_{\gamma'}}{c_{\gamma'}}\right|^2 + 4^{j+j'} |2^{-j}\lambda_* - 2^{-j'}\lambda'_*|^2$. Using that $\left|\tau_\gamma\right| \le C_1$ for some constant C_1 - independent of γ - cf. (1.10) and (2.1), we estimate

$$\begin{aligned}
\left|4^j \frac{\tau_\gamma}{c_\gamma} - 4^{j'} \frac{\tau_{\gamma'}}{c_{\gamma'}}\right| &= \left|4^j \frac{\tau_\gamma}{c_\gamma} - 4^{j'} \frac{\tau_{\gamma'}}{c_\gamma} + 4^{j'} \frac{\tau_{\gamma'}}{c_\gamma} - 4^{j'} \frac{\tau_{\gamma'}}{c_{\gamma'}}\right| \\
&\ge \left|4^j \frac{\tau_\gamma}{c_\gamma} - 4^{j'} \frac{\tau_{\gamma'}}{c_\gamma}\right| - \left|4^{j'} \frac{\tau_{\gamma'}}{c_\gamma} - 4^{j'} \frac{\tau_{\gamma'}}{c_{\gamma'}}\right| \\
&\ge C_{\mathrm{vel}}^{-1} \left|4^j \tau_\gamma - 4^{j'} \tau_{\gamma'}\right| - 4^{j'} C_1 \left|\frac{1}{c_\gamma} - \frac{1}{c_{\gamma'}}\right| \\
&= C_{\mathrm{vel}}^{-1} \left|4^j \tau_\gamma - 4^{j'} \tau_{\gamma'}\right| - 4^{j'} C_1 \left|\frac{1}{c(0, 2^{-j}\lambda_*)} - \frac{1}{c(0, 2^{-j'}\lambda'_*)}\right|.
\end{aligned}$$

Since the velocity c has (uniformly) bounded derivatives and is bounded below - cf. (1.10) - we conclude that $\left|\frac{1}{c(0,2^{-j}\lambda_*)} - \frac{1}{c(0,2^{-j'}\lambda'_*)}\right| \lesssim \left|2^{-j}\lambda_* - 2^{-j'}\lambda'_*\right|$. Consequently,

$$\left|4^j \tau_\gamma - 4^{j'} \tau_{\gamma'}\right|^2 \lesssim \left|4^j \frac{\tau_\gamma}{c_\gamma} - 4^{j'} \frac{\tau_{\gamma'}}{c_{\gamma'}}\right|^2 + 4^{2j'} \left|2^{-j}\lambda_* - 2^{-j'}\lambda'_*\right|^2.$$

Similarly, $\left|4^j \tau_\gamma - 4^{j'} \tau_{\gamma'}\right|^2 \lesssim \left|4^j \frac{\tau_\gamma}{c_\gamma} - 4^{j'} \frac{\tau_{\gamma'}}{c_{\gamma'}}\right|^2 + 4^{2j} \left|2^{-j}\lambda_* - 2^{-j'}\lambda'_*\right|^2$, and therefore

$$\begin{aligned}
\left|4^j \tau_\gamma - 4^{j'} \tau_{\gamma'}\right|^2 &\lesssim \left|4^j \frac{\tau_\gamma}{c_\gamma} - 4^{j'} \frac{\tau_{\gamma'}}{c_{\gamma'}}\right|^2 + \min\{4^{2j}, 4^{2j'}\} \left|2^{-j}\lambda_* - 2^{-j'}\lambda'_*\right|^2, \\
&\lesssim \left|4^j \frac{\tau_\gamma}{c_\gamma} - 4^{j'} \frac{\tau_{\gamma'}}{c_{\gamma'}}\right|^2 + 4^{j+j'} \left|2^{-j}\lambda_* - 2^{-j'}\lambda'_*\right|^2,
\end{aligned}$$

showing that the announced estimate indeed holds.

Step 2. *We show that* $L := \left|4^j \frac{\tau_\gamma}{c_\gamma} - 4^{j'} \frac{\tau_{\gamma'}}{c_{\gamma'}}\right|^2 \lesssim |4^j p_\gamma - 4^{j'} p_{\gamma'}|^2$.

We denote $\epsilon_\gamma = \frac{c_\gamma |p_{\gamma,*}|}{|\tau_\gamma|}$. Hence, by (8.7), $p_{\gamma,1} = \frac{|\tau_\gamma|}{c_\gamma}\sqrt{1-\epsilon_\gamma^2}$. By the grazing ray condition (4.6), $0 \le \epsilon_\gamma < 1$. We note that for $\gamma, \gamma' \in \Gamma_h^+$, $\tau_\gamma < 0$ and $\tau_{\gamma'} < 0$, while $p_{\gamma,1} > 0$ and $p_{\gamma',1} > 0$. Keeping these facts in mind and using inequality: $0 \le \left(1-x^2\right)^{\frac{1}{2}} \left(1-y^2\right)^{\frac{1}{2}} + xy \le 1$, for $x, y \in [0,1] \times [0,1]$, we estimate

$$\begin{aligned}
L &= 4^{2j} \frac{\tau_\gamma^2}{c_\gamma^2} + 4^{2j'} \frac{\tau_{\gamma'}^2}{c_{\gamma'}^2} - 2 \cdot 4^{j+j'} \frac{|\tau_\gamma||\tau_{\gamma'}|}{c_\gamma c_{\gamma'}} \\
&\le 4^{2j} \frac{\tau_\gamma^2}{c_\gamma^2} + 4^{2j'} \frac{\tau_{\gamma'}^2}{c_{\gamma'}^2} - 2 \cdot 4^{j+j'} \frac{|\tau_\gamma||\tau_{\gamma'}|}{c_\gamma c_{\gamma'}} \left[\left(1-\epsilon_\gamma^2\right)^{\frac{1}{2}} \left(1-\epsilon_{\gamma'}^2\right)^{\frac{1}{2}} + \epsilon_\gamma \epsilon_{\gamma'}\right] \\
&= 4^{2j} \frac{\tau_\gamma^2}{c_\gamma^2} + 4^{2j'} \frac{\tau_{\gamma'}^2}{c_{\gamma'}^2} - 2 \cdot 4^{j+j'} \frac{|\tau_\gamma||\tau_{\gamma'}|}{c_\gamma c_{\gamma'}} \left(1-\epsilon_\gamma^2\right)^{\frac{1}{2}} \left(1-\epsilon_{\gamma'}^2\right)^{\frac{1}{2}} - 2 \cdot 4^{j+j'} |p_{\gamma,*}||p_{\gamma',*}| \\
&= 4^{2j} \frac{\tau_\gamma^2}{c_\gamma^2} + 4^{2j'} \frac{\tau_{\gamma'}^2}{c_{\gamma'}^2} - 2 \cdot 4^{j+j'} \big(\tfrac{\tau_\gamma^2}{c_\gamma^2} - |p_{\gamma,*}|^2\big)^{\frac{1}{2}} \big(\tfrac{\tau_{\gamma'}^2}{c_{\gamma'}^2} - |p_{\gamma',*}|^2\big)^{\frac{1}{2}} - 2 \cdot 4^{j+j'} |p_{\gamma,*}||p_{\gamma',*}|.
\end{aligned}$$

Using the arithmetic-geometric means inequality: $4^{2j} |p_{\gamma,*}|^2 + 4^{2j'} |p_{\gamma',*}|^2 \le 2 \cdot 4^{j+j'} |p_{\gamma,*}| \times |p_{\gamma',*}|$ we conclude that

$$\begin{aligned}
L &\le 4^{2j} \big(\tfrac{\tau_\gamma^2}{c_\gamma^2} - |p_{\gamma,*}|^2\big) + 4^{2j'} \big(\tfrac{\tau_{\gamma'}^2}{c_{\gamma'}^2} - |p_{\gamma',*}|^2\big) - 2 \cdot 4^{j+j'} \big(\tfrac{\tau_\gamma^2}{c_\gamma^2} - |p_{\gamma,*}|^2\big)^{\frac{1}{2}} \big(\tfrac{\tau_{\gamma'}^2}{c_{\gamma'}^2} - |p_{\gamma',*}|^2\big)^{\frac{1}{2}} \\
&= 4^{2j} p_{\gamma,1}^2 + 4^{2j'} p_{\gamma',1}^2 - 2 \cdot 4^{j+j'} \cdot p_{\gamma,1} \cdot p_{\gamma',1} = \left(4^j p_{\gamma,1} - 4^{j'} p_{\gamma',1}\right)^2 \le |4^j p_\gamma - 4^{j'} p_{\gamma'}|^2,
\end{aligned}$$

as claimed.

Step 3. *Finally, we combine Steps 1 and 2 to deduce* (8.8). The proof for Γ_h^- is similar, with the difference that a minus signs is present in (8.7). □

We may now prove the announced result.

Proof of Theorem 6.5. We consider one of the families, $\Upsilon^{h,+}$ or $\Upsilon^{h,-}$, and drop the superscript $+$. We verify the conditions in Definition 3.1.

Step 1. *Estimates for* ω_γ^h *and* ξ_γ^h.

By definition, $\omega_\gamma^h = 4^j$ - cf. (6.18). Moreover, using (6.11), the fact that c is bounded below, and the non-tangential propagation estimate in (4.7) we conclude that $|p_\gamma(t_\gamma)| \asymp 1$. Using the fact that the Hamiltonian is constant on its flow, we can propagate this estimate to $t = 0$:

$$1 \asymp |p_\gamma(t_\gamma)| \asymp |c(x(t_\gamma)| |p_\gamma(t_\gamma)| = |H(x(t_\gamma), p(t_\gamma))| = |H(x(0), p(0))| \asymp |p_\gamma(0)|.$$

Hence $|\xi_\gamma^h| = \frac{4^j}{2\pi} |p_\gamma(0)| \asymp 4^j = \omega_\gamma^h$. This establishes one of the properties that we need in order to check the well-spreadness of $\Gamma_h^\pm$, and, additionally, it allows us to invoke Lemma 2.6 for this family of parameters.

Since $\omega_\gamma^h = \omega_\gamma^{st} = 4^j$, in what follows we write unambiguously ω_γ.

Step 2. *Some constants*. Recall the assumption in (4.5). Let us Taylor expand:

$$x_\gamma(t) = x_\gamma(t_\gamma) + (t - t_\gamma)y_\gamma(t), \tag{8.9}$$

where $y_{\gamma,i} \in C([-T, T])$ uniformly on γ by Lemma 2.6. Since $x_{\gamma,1}(t_\gamma) = 0$, Claim 6.1 allows us to invoke the *cone condition* in (4.3) and deduce that

$$\left|y_{\gamma,1}(t)\right| \geq \delta > 0, \qquad t \in [-T, T],$$

for some constant $\delta > 0$. In addition, by Lemma 2.6,

$$C_1 := \sup_{\gamma\in\Gamma_h} \sup_{t\in[0,T]} \max\left\{\left|y_\gamma(t)\right|, \left|\dot{p}_\gamma(t)\right|\right\} \tag{8.10}$$

is finite. We let $\varepsilon := \frac{1}{2}\min\{\frac{1}{2}\frac{\delta}{C_1}, \delta\}$ and note that

$$\left|y_{\gamma,1}(t)\right| \geq \varepsilon\left(\left|y_{\gamma,*}(t)\right| + 1\right), \qquad t \in [-T, T]. \tag{8.11}$$

Step 3. *We show that* $|a^{st}_\gamma - a^{st}_{\gamma'}| \lesssim |a^h_\gamma - a^h_{\gamma'}| + 1$.

According to the definitions,

$$|a^h_\gamma - a^h_{\gamma'}| = |x_\gamma(0) - x_{\gamma'}(0)|,$$
$$|a^{st}_\gamma - a^{st}_{\gamma'}| = |2^{-j}\lambda - 2^{-j'}\lambda'|.$$

By Lemma 2.6

$$|x_\gamma(t_\gamma) - x_{\gamma'}(t_\gamma)|^2 \lesssim |x_\gamma(0) - x_{\gamma'}(0)|^2 + 1.$$

Hence, it suffices to show that

$$|2^{-j}\lambda - 2^{-j'}\lambda'|^2 \lesssim |x_\gamma(t_\gamma) - x_{\gamma'}(t_\gamma)|^2. \tag{8.12}$$

To this end, we use the linearization in (8.9),

$$x_{\gamma'}(t_\gamma) = x_{\gamma'}(t_{\gamma'}) + (t_\gamma - t_{\gamma'})y_{\gamma'}(t),$$

and write

$$|x_\gamma(t_\gamma) - x_{\gamma'}(t_\gamma)| = |x_\gamma(t_\gamma) - x_{\gamma'}(t_{\gamma'}) - (t_\gamma - t_{\gamma'})y_{\gamma'}(t_\gamma)|.$$

Recall that $t_\gamma = 2^{-j}\lambda_1$ and $x_\gamma(t_\gamma) = (0, 2^{-j}\lambda_*)$. We use (8.11) to estimate

$$\begin{aligned}
|x_\gamma(t_\gamma) - x_{\gamma'}(t_\gamma)| &= \left|(-(t_\gamma - t_{\gamma'})y_{\gamma',1}(t_\gamma), x_{\gamma,*}(t_\gamma) - x_{\gamma',*}(t_{\gamma'}) - (t_\gamma - t_{\gamma'})y_{\gamma',*}(t_\gamma)\right| \\
&\asymp \left|y_{\gamma',1}(t_\gamma)\right| \left|2^{-j}\lambda_1 - 2^{-j'}\lambda'_1\right| + \left|2^{-j}\lambda_* - 2^{-j'}\lambda'_* + (2^{-j}\lambda_1 - 2^{-j'}\lambda'_1)y_{\gamma',*}(t_\gamma)\right| \\
&\asymp \left|y_{\gamma',1}(t_\gamma)\right| \left|2^{-j}\lambda_1 - 2^{-j'}\lambda'_1\right| + \varepsilon\left|2^{-j}\lambda_* - 2^{-j'}\lambda'_* + y_{\gamma',*}(t_\gamma)(2^{-j}\lambda_1 - 2^{-j'}\lambda'_1)\right|
\end{aligned}$$

$$\geq \varepsilon|2^{-j}\lambda_* - 2^{-j'}\lambda'_*| + \left(\left|y_{\gamma',1}(t_\gamma)\right| - \varepsilon\left|y_{\gamma',*}(t_\gamma)\right|\right)|2^{-j}\lambda_1 - 2^{-j'}\lambda'_1|$$
$$\geq \varepsilon|2^{-j}\lambda_* - 2^{-j'}\lambda'_*| + \varepsilon|2^{-j}\lambda_1 - 2^{-j'}\lambda'_1| \asymp |2^{-j}\lambda - 2^{-j'}\lambda'|.$$

Hence, (8.12) follows.

Step 4. *Estimates for* $\mathcal{M}^h_\gamma$.

By Claim 6.2, we know that $M_\gamma(t_\gamma)$ is symmetric, $\|M_\gamma(t_\gamma)\| \lesssim 1$ and $\Im(M_\gamma(t_\gamma)) \gtrsim I_d$. Those conclusions extend to $M_\gamma(0)$, since propagation preserves these conditions with different time dependent constants. This is stated in Lemma 2.6 for forward propagation $0 \mapsto t$, but the same conclusion is valid with an arbitrary initial time. (The general reference for this fact is [22, Lemma 2.56].) Since $\mathcal{M}^h_\gamma = (2\pi)^{-1}M_\gamma(0)$, the conclusion follows.

Step 5. *Estimates for* $\mathcal{A}^h_\gamma$.

By definition, $A_\gamma(t_\gamma) = 2^{\frac{d}{4}}4^{j\frac{d}{4}}$ - cf. (6.17). Since, by Step 4, $\|M_\gamma(0)\| \lesssim 1$ and $|t_\gamma| \leq C_{h,\sup}$, by Lemma 2.6 we conclude that $|A_\gamma(0)| \asymp |A_\gamma(t_\gamma)| \asymp 4^{j\frac{d}{4}}$. Hence, the choice made in (6.18) yields $\left|\mathcal{A}^h_\gamma\right| = 4^{-j\frac{d}{4}}A_\gamma(0) \asymp 1$, as desired.

Step 6. *We show that*

$$\left|\omega_\gamma p_\gamma(t_\gamma) - \omega_{\gamma'}p_{\gamma'}(t_{\gamma'})\right|^2 \lesssim \left|\omega_\gamma p_\gamma(t_\gamma) - \omega_{\gamma'}p_{\gamma'}(t_\gamma)\right|^2 + \omega_\gamma\omega_{\gamma'}|2^{-j}\lambda_1 - 2^{-j'}\lambda'_1|^2. \quad (8.13)$$

To see this, we assume without loss of generality that $\omega_\gamma \geq \omega_{\gamma'}$ and use the mean value theorem to find points $\bar{t}_{\gamma',i} \in [0, C_{h,\sup}]$ such that

$$\begin{aligned}
&\left|\omega_\gamma p_\gamma(t_\gamma) - \omega_{\gamma'}p_{\gamma'}(t_\gamma)\right| \\
&\quad \geq \left|\omega_\gamma p_\gamma(t_\gamma) - \omega_{\gamma'}p_{\gamma'}(t_{\gamma'})\right| - \omega_{\gamma'}\left(\sum_{i=1}^d \left|\dot{p}_{\gamma',i}(\bar{t}_{\gamma',i})\right|\right)\left|2^{-j}\lambda_1 - 2^{-j'}\lambda'_1\right| \\
&\quad \geq \left|\omega_\gamma p_\gamma(t_\gamma) - \omega_{\gamma'}p_{\gamma'}(t_{\gamma'})\right| - C_1 \cdot \sqrt{\omega_\gamma}\sqrt{\omega_{\gamma'}}|2^{-j}\lambda_1 - 2^{-j'}\lambda'_1|,
\end{aligned}$$

where C_1 is given by (8.10). Therefore, (8.13) follows.

Step 7. *We show that*

$$\left|\xi_{j,k} - \xi_{j',k'}\right|^2 \lesssim \left|\omega_\gamma p_\gamma(t_\gamma) - \omega_{\gamma'}p_{\gamma'}(t_{\gamma'})\right|^2 + \omega_\gamma\omega_{\gamma'}|2^{-j}\lambda_* - 2^{-j'}\lambda'_*|^2.$$

Since $\omega_\gamma = 4^j$ - cf. (6.18), this is just the content of Lemma 8.1.

Step 8. *We show that* $d((a^h_\gamma, \xi^h_\gamma), (a^h_{\gamma'}, \xi^h_{\gamma'})) \gtrsim d((a^{st}_\gamma, \xi^{st}_\gamma), (a^{st}_{\gamma'}, \xi^{st}_{\gamma'}))$, $\gamma, \gamma' \in \Gamma^\pm_h$.

We combine the previous steps and Lemma 2.6 to obtain:

$$\begin{aligned}
d((a^{st}_\gamma, \xi^{st}_\gamma), (a^{st}_{\gamma'}, \xi^{st}_{\gamma'})) &\asymp \omega_\gamma\omega_{\gamma'}|2^{-j}\lambda - 2^{-j'}\lambda'|^2 + \left|\xi_{j,k} - \xi_{j',k'}\right|^2 \\
&\lesssim \left|\omega_\gamma p_\gamma(t_\gamma) - \omega_{\gamma'}p_{\gamma'}(t_{\gamma'})\right|^2 + \omega_\gamma\omega_{\gamma'}|2^{-j}\lambda - 2^{-j'}\lambda'|^2 && \text{by Step 7} \\
&\lesssim \left|\omega_\gamma p_\gamma(t_\gamma) - \omega_{\gamma'}p_{\gamma'}(t_\gamma)\right|^2 + \omega_\gamma\omega_{\gamma'}|2^{-j}\lambda - 2^{-j'}\lambda'|^2 && \text{by Step 6} \\
&\lesssim \left|\omega_\gamma p_\gamma(t_\gamma) - \omega_{\gamma'}p_{\gamma'}(t_\gamma)\right|^2 + \omega_\gamma\omega_{\gamma'}|x_\gamma(t_\gamma) - x_{\gamma'}(t_\gamma)|^2 && \text{by (8.12)}
\end{aligned}$$

$$\asymp d\big(\big(x_\gamma(t_\gamma), \omega_\gamma p_\gamma(t_\gamma)\big), \big(x_{\gamma'}(t_\gamma), \omega_{\gamma'} p_{\gamma'}(t_\gamma)\big)\big) \qquad \text{by (2.17)}$$
$$\asymp d((a^h_\gamma, \xi^h_\gamma), (a^h_{\gamma'}, \xi^h_{\gamma'})). \qquad \text{by (2.18)}$$

This concludes the proof. □

9. Table of notation

Symbol	Description	Ref.
$\mathbb{R}^d_+$	$\mathbb{R}^d_+ = (0, +\infty) \times \mathbb{R}^{d-1}$.	Sec. 1.5
$\mathbb{R}^d_T$	$\mathbb{R}^d_T = [-T, T] \times \mathbb{R}^{d-1}$.	Sec. 1.5
x_*	$x = (x_1, x_*) \in \mathbb{R}^d$, with $x_1 \in \mathbb{R}$ and $x_* \in \mathbb{R}^{d-1}$	Sec. 1.5
$c = c(x)$	Velocity function.	
$h = h(t, x_*)$	Boundary data for the Dirichlet problem.	Sec. 4.1.1
$[C_{h,\inf}, C_{h,\sup}]$	Temporal support of the boundary data.	Sec. 4.1.1
C_{graz}	Constant related to the no-grazing ray assumption.	Sec. 4.1.1
$H^\pm = H^\pm(x, p)$	Signed Hamiltonian functions.	(2.8)
$H = H(x, p)$	Denotes generically either H^+ or H^-.	
$\sigma(x, D)$	Kohn-Nirenberg quantization of a symbol σ	Sec. 1.5
$\xi_{j,k}$	Center for the frequency cover.	Sec. 2.1
$\tilde{\xi}_{j,k}$	Approximately normalized version of $\xi_{j,k}$.	(2.1)
Λ	A lattice within $\mathbb{R}^d$. Throughout most of the text, the choice of Λ is fixed by Theorem 2.1	
Γ	Basic scale-angle-position index set.	(2.3)
Γ_*	Superset of Γ augmented with zero-scale.	(2.4)
Γ_0	A generic subset of Γ.	
Γ_h	A subset of Γ related to the frame expansion of h.	Sec. 4.2
$\Gamma^\pm_h$	Two subsets Γ^+_h and Γ^-_h that partition Γ_h.	(4.8)
γ	Generic element of (a subset of) Γ_*. We refer implicitly to the notation $\gamma = (j, k, \lambda)$.	Sec. 2.1
$\mathcal{S} = \mathcal{S}_\gamma$	A function that maps an index γ into a tuple of initial conditions for a GB.	Sec. 2.4
$\mathcal{S}^{st} = \mathcal{S}^{st}_\gamma$	The standard choice for such a map.	Sec. 2.5
$\mathcal{S}^{h,\pm} = \mathcal{S}^{h,\pm}_\gamma$	Two particular maps defined on $\Gamma^\pm_h$ respectively, constructed in terms of the boundary value h.	Sec. 6
$\Phi^\pm_\gamma$	GB associated with γ by means of an implicit map $\mathcal{S}_\gamma$.	Sec. 2.4
Φ_γ	Denotes generically either Φ^+_γ or Φ^+_γ.	
$\Phi^{st,\pm}_\gamma$	The beams associated with $\mathcal{S}^{st}_\gamma$.	Sec. 2.5
$\Phi^{h,\pm}_\gamma$	The beam associated with $\mathcal{S}^{h,\pm}_\gamma$. Here the mode is *determined* by whether $\gamma \in \Gamma^+_h$ or Γ^-_h.	Sec. 6.1
φ	Normalized Gaussian function.	(2.2)
φ_γ	Frame element associated with $\gamma \in \Gamma_*$.	Sec. 2.1
Υ	Generic set of GB parameters, indexed by a corresponding function $\mathcal{S}$.	Remark 2.4
Υ^{st}	The standard choice for such a set.	Sec. 2.5
$\Upsilon^{h,\pm}$	Two particular such sets associated with h.	Sec. 6.1
$F = \mathbb{O}^m(I, \Upsilon)$	A family of functions F_γ that vanishes to oder m on the centers of the beams defined by Υ, uniformly on the time interval I.	Definition 3.3
$F = \mathbb{O}^m_\geq(I, \Upsilon)$	Functions with vanishing order at least m.	Definition 3.4

Acknowledgments

M.B. gratefully acknowledges support from the Fondazione CRT. M.V.d.H. gratefully acknowledges support from the Simons Foundation under the MATH + X program, the National Science Foundation under grant DMS-1559587, and the corporate members of the Geo-Mathematical Group at Rice University. J.L.R. gratefully acknowledges support from the Austrian Science Fund (FWF): P 29462 - N35 and Y 1199; from a Marie Curie fellowship, within the 7th European Community Framework program, under grant PIIF-GA-2012-327063; and from the WWTF grant INSIGHT (MA16-053).

References

[1] A. Aldroubi, C. Cabrelli, U.M. Molter, Wavelets on irregular grids with arbitrary dilation matrices and frame atoms for $L^2(\mathbb{R}^d)$, Appl. Comput. Harmon. Anal. 17 (2) (2004) 119–140.

[2] F. Andersson, M. Carlsson, L. Tenorio, On the representation of functions with Gaussian wave packets, J. Fourier Anal. Appl. 18 (1) (2012) 146–181.

[3] V.M. Babich, V.S. Buldyrev, Asimptoticheskie metody v zadachakh difraktsii korotkikh voln. Tom I, Izdat. “Nauka”, Moscow, 1972. Metod ètalonnykh zadach. [The method of canonical problems], with the collaboration of M.M. Popov and I.A. Molotkov.

[4] G. Bao, J. Qian, L. Ying, H. Zhang, A convergent multiscale Gaussian-beam parametrix for the wave equation, Commun. Partial Differ. Equ. 38 (1) (2013) 92–134.

[5] M. Berra, M.V. de Hoop, J.L. Romero, A multi-scale Gaussian beam parametrix for the wave equation: the Dirichlet boundary value problem, Technical report. arXiv:1705.00337.

[6] N. Bleistein, S.H. Gray, Amplitude calculations for 3D Gaussian beam migration using complex-valued traveltimes, Inverse Probl. 26 (8) (2010) 085017.

[7] E.J. Candès, L. Demanet, The curvelet representation of wave propagators is optimally sparse, Commun. Pure Appl. Math. 58 (11) (2005) 1472–1528.

[8] V. Červenỳ, M.M. Popov, I. Pšenčík, Computation of wave fields in inhomogeneous media—Gaussian beam approach, Geophys. J. Int. 70 (1) (1982) 109–128.

[9] J. Chandra, P.W. Davis, Linear generalizations of Gronwall’s inequality, Proc. Am. Math. Soc. 60 (1976) 157–160 (1977).

[10] A. Córdoba, C. Fefferman, Wave packets and Fourier integral operators, Commun. Partial Differ. Equ. 3 (11) (1978) 979–1005.

[11] M.V. de Hoop, K. Gröchenig, J.L. Romero, Exact and approximate expansions with pure Gaussian wave packets, SIAM J. Math. Anal. 46 (3) (2014) 2229–2253.

[12] M.V. de Hoop, S.F. Holman, H.F. Smith, G. Uhlmann, Regularity and multi-scale discretization of the solution construction of hyperbolic evolution equations with limited smoothness, Appl. Comput. Harmon. Anal. 33 (3) (2012) 330–353.

[13] M.V. de Hoop, H. Smith, G. Uhlmann, R.D. van der Hilst, Seismic imaging with the generalized Radon transform: a curvelet transform perspective, Inverse Probl. 25 (2) (2009) 025005.

[14] M.V. de Hoop, G. Uhlmann, A. Vasy, H. Wendt, Multiscale discrete approximations of Fourier integral operators associated with canonical transformations and caustics, Multiscale Model. Simul. 11 (2) (2013) 566–585.

[15] L. Demanet, L. Ying, Wave atoms and sparsity of oscillatory patterns, Appl. Comput. Harmon. Anal. 23 (3) (2007) 368–387.

[16] L. Demanet, L. Ying, Wave atoms and time upscaling of wave equations, Numer. Math. 113 (1) (2009) 1–71.

[17] G.A. Deschamps, Gaussian beam as a bundle of complex rays, Electron. Lett. 7 (23) (November 1971) 684–685.

[18] D.-A. Geba, D. Tataru, A phase space transform adapted to the wave equation, Commun. Partial Differ. Equ. 32 (7–9) (2007) 1065–1101.

[19] N.R. Hill, Gaussian beam migration, Geophysics 55 (11) (1990) 1416–1428.

[20] L. Hörmander, On the Existence and the Regularity of Solutions of Linear Pseudo-Differential Equations, Série des Conférences de l’Union Mathématique Internationale, vol. 1, Secrétariat de l’Enseignement Mathématique, Université de Genève, Geneva, 1971, Monographie No. 18 de l’Enseignement Mathématique.

[21] A. Kachalov, M. Popov, Application of the method of summation of Gaussian beams for calculation of high-frequency wave fields, Sov. Phys. Dokl. 26 (1981) 604.

[22] A. Katchalov, Y. Kurylev, M. Lassas, Inverse Boundary Spectral Problems, Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics, vol. 123, Chapman & Hall/CRC, Boca Raton, FL, 2001.

[23] J.B. Keller, W. Streifer, Complex rays with an application to Gaussian beams, J. Opt. Soc. Am. 61 (1) (Jan 1971) 40–43.

[24] H. Koch, D. Tataru, Dispersive estimates for principally normal pseudodifferential operators, Commun. Pure Appl. Math. 58 (2) (2005) 217–284.

[25] A. Laptev, I.M. Sigal, Global Fourier integral operators and semiclassical asymptotics, Rev. Math. Phys. 12 (5) (2000) 749–766.

[26] I. Lasiecka, J.-L. Lions, R. Triggiani, Nonhomogeneous boundary value problems for second order hyperbolic operators, J. Math. Pures Appl. (9) 65 (2) (1986) 149–192.

[27] J.-L. Lions, E. Magenes, Non-Homogeneous Boundary Value Problems and Applications. Vol. I, Die Grundlehren der mathematischen Wissenschaften, band 181, Springer-Verlag, New York-Heidelberg, 1972, translated from the French by P. Kenneth.

[28] H. Liu, O. Runborg, N.M. Tanushev, Error estimates for Gaussian beam superpositions, Math. Comput. 82 (282) (2013) 919–952.

[29] J. Lu, X. Yang, Frozen Gaussian approximation for high frequency wave propagation, Commun. Math. Sci. 9 (3) (2011) 663–683.

[30] J. Lu, X. Yang, Convergence of frozen Gaussian approximation for high-frequency wave propagation, Commun. Pure Appl. Math. 65 (6) (2012) 759–789.

[31] R.B. Melrose, Local Fourier-Airy integral operators, Duke Math. J. 42 (4) (1975) 583–604.

[32] R.B. Melrose, Microlocal parametrices for diffractive boundary value problems, Duke Math. J. 42 (4) (1975) 605–635.

[33] R.B. Melrose, M. Taylor, Boundary problems for wave equations with grazing and gliding rays, Manuscript, retrieved from: www.unc.edu/math/faculty/met/glide.pdf on 02/20/2017.

[34] M. Motamed, O. Runborg, Taylor expansion and discretization errors in Gaussian beam superposition, Wave Motion 47 (7) (2010) 421–439.

[35] M. Nielsen, Frames for decomposition spaces generated by a single function, Collect. Math. 65 (2) (2014) 183–201.

[36] L. Nirenberg, Lectures on linear partial differential equations, Usp. Mat. Nauk 30 (4(184)) (1975) 147–204, translated from the English by Ju. V. Egorov (Amer. Math. Soc., Providence, R.I., 1973).

[37] T.J.P.M. Op 't Root, C.C. Stolk, M.V. de Hoop, Linearized inverse scattering based on seismic reverse time migration, J. Math. Pures Appl. (9) 98 (2) (2012) 211–238.

[38] M.M. Popov, A new method for calculating wave fields in high-frequency approximation, in: Mathematical Questions in the Theory of Wave Propagation, 11, Zap. Nauchn. Semin. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) 104 (1981) 195–216, 239–240.

[39] M.M. Popov, A new method of computation of wave fields using Gaussian beams, Wave Motion 4 (1) (1982) 85–97.

[40] J. Qian, L. Ying, Fast Gaussian wavepacket transforms and Gaussian beams for the Schrödinger equation, J. Comput. Phys. 229 (20) (2010) 7848–7873.

[41] J. Qian, L. Ying, Fast multiscale Gaussian wavepacket transforms and multiscale Gaussian beams for the wave equation, Multiscale Model. Simul. 8 (5) (2010) 1803–1837.

[42] J. Ralston, Gaussian beams and the propagation of singularities, in: Studies in Partial Differential Equations, in: MAA Stud. Math., vol. 23, Math. Assoc. America, Washington, DC, 1982, pp. 206–248.

[43] H.F. Smith, A Hardy space for Fourier integral operators, J. Geom. Anal. 8 (4) (1998) 629–653.

[44] H.F. Smith, A parametrix construction for wave equations with $C^{1,1}$ coefficients, Ann. Inst. Fourier (Grenoble) 48 (3) (1998) 797–835.

[45] C.C. Stolk, On the modeling and inversion of seismic data, Ph. D. thesis, Utrecht University, 2000.

[46] C.C. Stolk, A pseudodifferential equation with damping for one-way wave propagation in inhomogeneous acoustic media, Wave Motion 40 (2) (2004) 111–121.

[47] N.M. Tanushev, B. Engquist, R. Tsai, Gaussian beam decomposition of high frequency wave fields, J. Comput. Phys. 228 (23) (2009) 8856–8871.

[48] D. Tataru, Phase space transforms and microlocal analysis, in: Phase Space Analysis of Partial Differential Equations, vol. II, in: Pubbl. Cent. Ric. Mat. Ennio Giorgi, Scuola Norm. Sup., Pisa, 2004, pp. 505–524.

[49] A. Waters, A parametrix construction for the wave equation with low regularity coefficients using a frame of Gaussians, Commun. Math. Sci. 9 (1) (2011) 225–254.

[50] V.H. Weston, Wave splitting and the reflection operator for the wave equation, in: Inverse Methods in Action, Montpellier, 1989, in: Inverse Probl. Theoret. Imaging, Springer, Berlin, 1990, pp. 227–238.