MULTISCALE GAUSSIAN-BEAM METHOD FOR HIGH-FREQUENCY WAVE PROPAGATION AND INVERSE PROBLEMS

By

Chao Song

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Applied Mathematics — Doctor of Philosophy

2018

ABSTRACT

MULTISCALE GAUSSIAN-BEAM METHOD FOR HIGH-FREQUENCY WAVE PROPAGATION AND INVERSE PROBLEMS

By

Chao Song

The existence of Gaussian beam solution to hyperbolic PDEs has been known to the pure mathematics community since sometime in the 1960s [3]. It enjoys popularity afterwards due to its ability to resolve the caustics problem and its efficiency [49, 28, 31]. In this thesis, we will focus on the extension of the multi-scale Gaussian beam method and its application to seismic wave modeling and inversion.

In the first part of thesis, we discuss the application of the multi-scale Gaussian beam method to the inverse problem. A new multi-scale Gaussian beam method is introduced for carrying out true-amplitude prestack migration of acoustic waves. After applying the Born approximation, the migration process is considered as shooting two beams simultaneously from the subsurface point which we want to image. The Multi-scale Gaussian Wavepacket transform provides an efficient and accurate way for both decomposing the perturbation field and initializing Gaussian beam solution. Moreover, we can prescribe both the region of imaging and the range of dipping angles by shooting beams from a subsurface point in the region of imaging. We prove the imaging condition equation rigorously and conduct error analysis. Some numerical approximations are derived to improve the efficiency further. Numerical results in the two-dimensional space demonstrate the performance of the proposed migration algorithm.

In the second part of thesis, we propose a new multiscale Gaussian beam method with reinitialization to solve the elastic wave equation in the high frequency regime with different boundary conditions. A novel multiscale transform is proposed to decompose any arbitrary vector-valued function to multiple Gaussian wavepackets with various resolution. After the step of initializing, we derive various rules corresponding to different types of reflection cases. To improve the efficiency and accuracy, we develop a new reinitialization strategy based on the stationary phase approximation method to sharpen each single beam ansatz. This is especially useful and necessary in some reflection cases. Numerical examples with various parameters demonstrate the correctness and robustness of the whole method. There are two boundary conditions considered here, the periodic and the Dirichlet boundary condition. In the end, we show that the convergence rate of the proposed multiscale Gaussian beam method follows the convergence rate of the classical Gaussian beam solution, i.e. $O\left(\frac{1}{\sqrt{\omega}}\right)$.

To my family

ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to my advisor Prof. Qian, Jianliang, for his continuous support of my Ph.D. study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor for my Ph.D. study.

I would like to express my gratitude to Prof. Zhou, Zhengfang, Prof. Cheng, Yingda and Prof. Tang, Moxun for serving on my thesis committee and critical comments and valuable feedbacks. I also would like to thank all the professors whom I took classes from at MSU, for their uniform excellence in teaching.

I would like to thank the support staff of Department of Mathematics for all their help with administrative matters. I would like to acknowledge my dear friends Ruochuan Zhang, Jun Du, Wangtao Lu and Wenbin Li for their friendship over the past six years. They made the experience of graduate school more rewarding.

Many thanks go to my parents for supporting me spiritually throughout writing this thesis and my life in general. Last but not the least, I would like to express my profound gratitude to my wife Yu, who has shown me the deepest love, kindness, and patience throughout our time together. This dissertation stands as a testament to your unconditional love and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ALGORITHMS

Chapter 1

Introduction

Hyperbolic PDEs arise in a variety of practical applications, ranging from acoustics, elasticity, electromagnetics to geophysics. Therefore, it is desirable to develop fast and efficient algorithms to solve this family of PDEs. Moreover, efficient algorithms are also desirable in a lot of inverse problems modeled by hyperbolic PDEs. The term 'high frequency' in the high frequency wave propagation is defined relative to the low-frequency background slowness in the model. Therefore, all medium appeared in this thesis is smooth, if not specified.

It is very costly for the direct method, the finite difference method or the finite element method for example, to simulate the high frequency wave propagation, since fine grid mesh is required by these methods to capture the oscillation. Therefore, some alternative methods have been developed, such as the traditional geometrical-optics method (WKBJ ansatz), which is required to solve a pair of equations. The first one, which is called eikonal equation, is a first-order nonlinear PDE. The second one is the transport equation relying on the differentiability of the result of eikonal equation. It yields faithful asymptotic solutions before caustics occur. However, the amplitude function governed by the transport equation breaks down around the caustics [6, 22, 36], where the phase function is multi-valued [36, 39]. The appearance of caustics is inevitable even in the smooth medium [54].

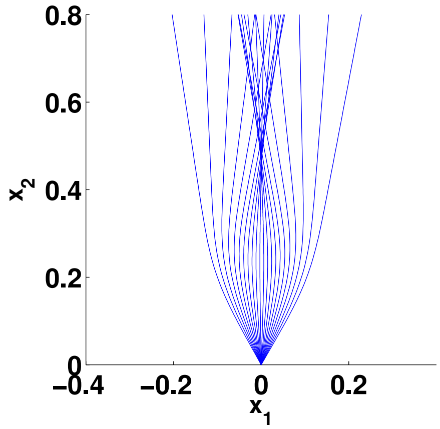


Figure 1.1: Caustics Appearance in Gaussian Velocity

One of the alternatives is the Gaussian beam method [49, 51, 55, 15]. To resolve the caustics problems automatically, the Gaussian beam method relaxes the restriction that the phase term is real-valued. The single Gaussian beam ansatz is made up of a complex-valued phase function and a complex-valued amplitude function near its central ray. Away from the ray path, the beam decays rapidly as a Gaussian profile. The superposition of many single beams will be an asymptotically correct solution of the hyperbolic PDE in the sense that both the initial condition and the boundary condition are satisfied asymptotically as well as the PDE itself.

There are two methods based on Gaussian beam method presented in this thesis to resolve two related problems. The first one is about the simulation of the elastic wave in a bounded domain. The second one is using a novel Gaussian beam method to reconstruct reflectors under the surface by the data received along the boundary.

1.1 Prestack Inversion Process

In the first part, we propose a novel inversion and imaging procedure for finding the parameter of the medium by solving the linearized inverse scatter problem. There is a wide range of applications including seismic exploration, medical imaging and underwater acoustics, etc.

There are many different migration methods existed to explore the geological structure. The very early one was using the one-way wave equation [16, 17] to recover the accurate travel time and locations of reflectors. However, the popular finite difference and finite element method require extremely refined grid mesh in the high frequency regime to prevent dispersion error. This is also found in the modeling process.

Kirchhoff method [7, 8, 28], which bases on the asymptotic method, is mostly investigated and widely applied to resolve the high frequency pattern. The Kirchhoff method and all other similar ray-based methods are using the asymptotic approximation (high frequency) of the Green's function for the acoustic wave equation.

$$G(x, x_0, \omega) = Ae^{i\omega\tau(x,x_0)}, \tag{1.1}$$

where $\tau(x, x_0)$ is the traveltime from the point x_0 to the point x.

There exist several other problems in the Kirchhoff migration, although it is efficient and flexible. The first obstacle is the presence of caustics points. It is not able to characterize the structure in the presence of caustics, since they rely on the first-arrival traveltime [18, 27] instead of multivalued traveltime. Its usefulness was questioned by

[25, 27, 43] since the first-arrival traveltime in complex media usually do not correspond to the most energetic traveltimes crucial for imaging complex structures.

To overcome this issue, we use the Gaussian beam migration method [3, 45, 14].

The first obstacle is the way to initialize each beam solution efficiently, in other words, how to characterize the wave propagator by beam solution in a sparse form. Another closely related issue is the way to describe the wavefront of the perturbation. The wavefront defines the singularities of a function not only in spatial space but also with respect to its Fourier transform at each point. It is naturally adaptable to the high frequency wave propagation which can be considered as the propagation of singularity. From this point of view, the way to generate the decomposition should consider both the optimal representation of the wave propagator and the sufficient condition to reconstruct the target perturbation media. In some sense, some tradeoff should be obtained.

The second obstacle is about the rigorous mathematical analysis of the imaging condition. In this paper, the spread loss has to be proved to be compensated for in our new imaging method. It is well known that the Gaussian beam solution has $O(1/\sqrt{\omega})$ convergence rate as an asymptotic solution. We will take advantage of this convergence rate and the fact that the parabolic scaling principle is preserved along the propagation to conduct a rigorous error analysis. Besides, we assume that a set of geodesics (rays) have a consistent direction. The assumption simply means that the overturn rays don't exist in our model and this assumption is natural in the practical applications.

Our method enjoys several advantages. First, the flexibility of the Kirchhoff migration is preserved and imaging without losing multi-arrival in the general slowness is possible.

Second, due to the Multi-scale Gaussian wavepacket transform and the parabolic scaling principle employed, the sparse representation of the wave propagator is obtained. This decomposition also makes the reinitialization feasible during the propagation so that we have more control over the width of beam solution. Third, there is some redundancy in the data in the view of our imaging condition. In practice, the error is unavoidable, then this redundancy can help us to reduce the impact brought by the noise since the average value is employed to cancel the variation caused. Fourth, our imaging condition is performed in the time domain to avoid the extra Fourier transform on the data set. This feature makes our algorithm more applicable considering the large size of the trace dataset in the real world. The last feature is that the wavefront set being characterized by the Gaussian wavepacket enables us to image the subsurface partially in the sense of controlling the range of dipping angles.

There are some other true amplitude migration methods different from the ray-based method described above. Zhang *et al.* [56, 57] developed true-amplitude common-shot migration for heterogeneous media. Again, only the geometric spreading loss is recovered in these works. Other types of compensations, for example transmission loss compensation, are discussed in [21].

1.2 Elastic Wave Equation

The elastic wave equation is a good model to describe the seismic wave propagation in a uniform whole space and it has been used widely in the seismology community in both inverse and modeling problem [1]. Similar to its simpler form, i.e. the acoustic wave equation, the elastic wave equation will propagate oscillations in space and time when the initial or boundary condition contains oscillation of the small wavelength. We will develop a multi-scale Gaussian beam method to simulate the elastic wave in a bounded domain.

The first problem to apply the Gaussian Beam method is how to decompose any general initial condition to the form suitable to the beam profile. It is resolved by the Multiscale Gaussian Wavepacket transform developed in [48, 5] for the acoustic wave equation and single-scale transform for the Schrodinger equation [47]. Since the Hamiltonian of the wave equation is homogeneous of order one, a Gaussian beam should satisfy the parabolic scaling principle at any given time if it is satisfied at the first place. The wavepacket satisfying the parabolic scaling principle is defined as the wavelength of the typical oscillation of this wavepacket being equal to the square of the width of this wavepacket.

To propagate each wavepacket, the dynamics system can be obtained in the typical way. In this paper, we extend this idea further to the decomposition of the vector-valued initial value and preserve all the optimal properties of the Multiscale Gaussian Wavepacket transform.

The second problem is to derive the reflection dynamics. Unlike the Cauchy problem, we have to consider the reflected beams. Most recently in [4], a numerical method has been proposed for the acoustic wave equation. Other discussions can be found in [9, 49]. There will be more complex situations concerned in the elastic wave model as there are two different types of wave modes in the process. The difference between different wave modes requires extra efforts to preserve the accuracy.

The third problem is how to reinitialize Gaussian beams efficiently. Gaussian beam solutions behave well around the caustics, however, not so well in the long-term wave propagation. See more details illustrated in [48, 32, 42].

Therefore, there have been various methods developed to control the width of beams. One method, for example, is to call the Multiscale Gaussian wavepacket algorithm repeatedly during the propagation [48, 4]. A global time T is set at prior such that beams will be summarized after propagating for T and then decompose the resulting temporary wavefield to a new system of wavepackets. This process will be repeated several times.

A new reinitialization method is proposed in this thesis which claims to have more freedom and remains to be the asymptotic solution to the elastic wave equation. Instead of decomposing the general wavefield after summation each time, we target on each single beam in this new reinitialization method. It will give us more freedom to choose which beams needed to be reinitialized rather than all of them. Moreover, with applying the reinitialization strategy to each single beam, there exists the explicit expression.

1.3 Related Work

The fact that the Gaussian beam ansatz can be used to solve the wave equation has been known to the pure mathematical community since sometime in the 1960s [3]. Then it is applied to simulate the propagation of the singularity [49, 30]. A single Gaussian beam ansatz is an asymptotic solution concentrated on a single ray curve. The critical point is to have a global solution to the Hessian of the phase function so that the transport equation is well-defined. Moreover, the fact that imaginary part of the Hessian remains to be S.P.D. leads to a well-localized solution. The localization is justified theoretically [49, 40, 52].

Considering its ability to resolve multi-valued phase function automatically, the Gaussian beam method was firstly introduced as a seismic imaging method by Hill in the form of the poststack [28] and then the prestack migration procedure [29]. The performance of the Gaussian beam migration is further tested by the common shot geometry [26]. Most recently, a purely Eulerian computational approach was proposed in [33] which improves the numerical method's efficiency and its application in the semiclassical quantum mechanics has been proposed in [31]. See [9, 42] for other recent works.

Besides the Gaussian beam method, there are several possible ways to construct global asymptotic solutions for the wave equation even in the presence of caustics. The first approach is based on Ludwigs uniform asymptotic expansions at caustics [35, 10] which requires that the caustic structure is given. The second approach is is based on the Maslov canonical operator theory [41], which requires to identify where the caustics are at prior.

There is some recent advance in resolving the multi-valued traveltime problem. A new method called fast Huygens sweeping method has been proposed in [37, 38] to solve the Helmholtz equations in the inhomogeneous media and then it is used to solve Schrodinger equation [34]. They take advantage of the fact that eikonal equation is well-defined around the source point and Huygens-Kirchhoff secondary source principle.

There have been some recent advance in the optimal representation of the wave propagator [52, 11, 50]. It is closely related to the Fourier integral operator representation of the hyperbolic system and the special proposition of its phase function. The multiscale Gaussian Wavepacket transform is developed [48] for the wave equation and the single scale transform for the Schrodinger equation [47] which is also the fundamental basis of our algorithm. This difference comes from the different Hamiltonian for these two equations. Other paper also apply the Gaussian beam method in their true amplitude migration [46, 2], however, they do not require the parabolic scaling principle for Gaussian beams as our prestack inversion method does.

The parabolic scaling principle provides the theoretical basis for our new reinitialization method as well as the proof of the correctness of the imaging operator in our new multi-scale inversion algorithm. Other methods using the similar idea can be found in [11, 50]. There are various types of such wavepackets, curvelet [13, 12] and wave atoms [19, 20] for example. However, the Multi-scale Gaussian beam method is different from these methods in that the single Gaussian wavepacket corresponds to the single Gaussian beam at final time T, while the curvelet frame does not have this one-to-one relationship.

On the other hand, the Gaussian wavepacket transform has been proved a stable and efficient decomposition of the arbitrary function [48], equivalently, the wavepacket is a good characterization of the wavefront set. The Gaussian window function or Gabor frame [24] are both well-localized in the phase space as the Fourier transform of a Gaussian profile function is again a Gaussian profile function. The size of the Gaussian window function in the phase space can be determined by the Heisenberg Uncertainty Principle.

1.4 Contents

The remainder of this thesis is organized as follows. In Chapter 2 we present a brief introduction of constructing and propagating a single Gaussian beam, which is the foundation to the following derivation. We then describe the original multi-scale Gaussian wavepacket transform [48] in Section 2.2.

In Chapter 3, we propose a new prestack inversion process based on the multi-scale Gaussian beam method. We then modify the Gaussian wavepacket transform in Section 3.1 to adapt to the imaging operator. We then develop the new imaging operator with the help of Gaussian wavepacket transform and the Gaussian beam functions. Based on this operator, we propose the main inversion algorithm in Section 3.2.4. The next part is devoted to proving the correctness of this new algorithm in Section 3.3. In Section 3.4, we discuss the fast method to calculate the imaging operator. In the last section of this chapter, we select several well-designed numerical examples to justify the correctness of our analysis and the approximations mentioned earlier.

In Chapter 4 we propose the Multiscale Gaussian beam method to solve the elastic wave equation with highly oscillated initial condition. We first extend the Multiscale Gaussian wavepacket transform to the vector-valued initial condition in Section 4.3 and develop a new propagating dynamics for each single beam. After proposing the decomposition scheme, the reflection dynamics for the homogeneous Dirichlet Boundary value is derived in Section 4.5. The difference among various types of reflection is analyzed in Section 4.7.2 and a new efficient reinitialization method is proposed to resolve the problem from S-wave reflection in Section 4.6 and Section 4.7.3. The reference solution to the elastic wave equation in general case is provided by the Finite-

Difference Time-Domain (FDTD) with staggered grid [53] and is justified in Section B. In Section 4.8, several numerical experiments are conducted to show the correctness and the convergence rate of our new multi-scale Gaussian beam method.

Chapter 2

Single Gaussian beam ansatz and Multiscale Gaussian wavepacket Transform

2.1 Single Gaussian beam ansatz

The Gaussian beam solution itself is an asymptotic solution of the acoustic wave equation even around caustics.

$$\frac{1}{c^2(x)}\partial_t^2 u(x,t) - \Delta u(x,t) = 0, \tag{2.1}$$

where x is the point coordinate in the space $\mathbb{R}^d$ and $c(x)$ is smooth, positive and bounded away from zero. Similar to the Geometric-Optics ansatz, the Gaussian beam also assumes that the solution follows the form,

$$u(x,t) = A(x,t)e^{i\omega\tau(x,t)}, \tag{2.2}$$

where ω is a large wavenumber, $\tau(x,t)$ is the phase function and $A(x,t)$ is the amplitude function. The asymptotic solutions means $u(x,t)$ (2.2) satisfies the wave equation (2.1) with small error when frequency ω is large. After inserting equation (2.2) into equation (2.1) and organize all terms according to the order of ω, there will be two equations obtained, which are eikonal and transport equations, governing $\tau(x,t)$ and $A(x,t)$ respectively. They come from the leading orders in inverse power of the frequency ω.

$$\tau_t^2(x,t) - c^2|\nabla\tau(x,t)|^2 = 0 \tag{2.3}$$

$$2A_t\tau_t - 2c^2\nabla A \cdot \nabla\tau + A(\tau_{tt} - c^2 trace(\tau_{xx})) = 0. \tag{2.4}$$

Phase function τ:

After factoring out equation (2.3), there are two branches generated,

$$\tau_t \pm c(x)|\nabla\tau| = 0. \tag{2.5}$$

Equation (2.6) is a Hamilton-Jacobi equation with the Hamiltonian $G^{\pm}(x,p) = \pm c(x)|p|$. We consider the generic situation for the eikonal equation,

$$\tau_t + G(x, \nabla\tau(x,t)) = 0. \tag{2.6}$$

We apply the method of the characteristics to solve the eikonal equation (2.6).

$$\begin{aligned}
\frac{dx}{dt} &= G_p(x(t), p(t)), \quad x\Big|_{t=0} = x_0; \\
\frac{dp}{dt} &= -G_x(x(t), p(t)), \quad p\Big|_{t=0} = p_0.
\end{aligned} \tag{2.7}$$

where we define the ray trajectory $\gamma = \{(x(t), p(t)) : t \geq 0\}$, whose initial point is (x_0, p_0) in the phase space. We have that the momentum $p(t) = \nabla\tau(x(t), t)$ along the ray.

To derive the dynamics of Hessian matrices, we first differentiate the eikonal equation (2.6) with respect to t and x:

$$\tau_{tx}(x,t) + G_x(x, \nabla\tau(x,t)) + \tau_{xx}(x,t)G_p(x(t), \nabla\tau(x,t)) = 0, \tag{2.8}$$

$$\tau_{tt}(x,t) + G_p(x(t), \nabla\tau(x,t)) \cdot \tau_{xt}(x,t) = 0, \tag{2.9}$$

Differentiating equation (4.13) with respect to x yields

$$\tau_{txx} + G_{xx} + \tau_{xx}G_{xp} + (G_{xp})^T\tau_{xx} + \tau_{xx}G_{pp}\tau_{xx} + \tau_{xxx}G_p = 0. \tag{2.10}$$

Therefore, the Hessian $M(t) = \nabla\nabla\tau(x(t), t)$ satisfies the following Riccati equation,

$$\frac{dM}{dt} + G_{xx} + MG_{xp} + G_{xp}^T M + MG_{pp}M = 0. \quad M\Big|_{t=0} = i\epsilon I \tag{2.11}$$

The size parameter ϵ will be given after introducing the Gaussian wavepacket transform.

One of the most significant differences between the Gaussian beam and other ray-ansatz methods is that beams' phase functions $\tau(x,t)$ are complex-valued. Complexifying the equation guarantees a well-defined Hessian and a well-defined transport equation as a result. This is not true in general case [49]. Furthermore, the positive definite imaginary part is always true throughout the propagation for smooth ray trajectories.

Lemma 2.1.1. *If the Hamiltonian G is smooth enough, then the Hessian M(t) along the ray path γ has a positive-definite imaginary part, provided that it initially does.*

Transport Equation $A(x,t)$:

With $(x(t), p(t), M(t))$ well-defined along the way, we can solve the transport equation (2.4). Taking advantage of the fact that

$$\frac{dA(x(t),t)}{dt} = A_t + \nabla A \cdot G_p(x(t), p(t)),$$

equation (2.4) is reduced to

$$\frac{dA}{dt} + \frac{A}{2G}(v^2 trace(M) - G_x \cdot G_p - G_p^T M G_p) = 0. \quad A\Big|_{t=0} = A_0 \tag{2.12}$$

A single beam solution is in the following form,

$$U_{p_0}^{x_0}(x,t) = A(x,t)e^{i\omega\tau(x,t)}, \tag{2.13}$$

and the phase function is approximated by applying Taylor expansion around the central trajectory $\{x(t) : t \geq 0\}$ at time t,

$$\tau(x,t) = p(t) \cdot (x - x(t)) + \frac{1}{2}(x - x(t))^T M(t)(x - x(t)), \tag{2.14}$$

and the amplitude is approximated by its value at trajectory γ at the same moment,

$$A(x,t) = A(x(t),t) = A(t). \tag{2.15}$$

2.2 Multiscale Gaussian Wavepacket Transform

After constructing a single beam solution, the next problem is how to set up the initial condition for the ODE system. The answer is Multiscale Gaussian wavepacket transform, which will be introduced briefly in this section. More details about this phase space decomposition can be found in paper [48] and its single scale application can be found in paper [47]. The wavepacket transform is applied to L_2 functions f in the $\mathbb{R}^d$ space.

We first partition the Fourier space $\mathbb{R}^d$ into several Cartesian coronae C_l for $l \geq 1$ as

$$C_l = \{\xi = (\xi_1, \xi_2, \cdots, \xi_d) : \max_{1 \leq i \leq d} |\xi_i| \in [4^{l-1}, 4^l]\}.$$

Now it is obvious to see that the L_2 norm of ξ in C_l is $O(4^l)$. For each C_l, we can further partition it into multiple windows with width 2^l,

$$B_{l,i} = \prod_{s=1}^{d} [2^l i_s, 2^l(i_s + 1)],$$

where the integer multiindex $(i_1, i_2, \cdots, i_d)$ is any possible choice such that the box is in the l^{th} layer, i.e. $B_{l,i} \subset C_l$. After defining these cell boxes $B_{l,i}$, we can define the Gaussian profile function $g_{l,i}$ associated with the box $B_{l,i}$ by the following formula,

$$g_{l,i}(\xi) \approx e^{-\left(\frac{|\xi - \xi_{l,i}|}{\sigma_l}\right)^2}, \tag{2.16}$$

where $\xi_{l,i}$ is the center of the box $B_{l,i}$ and $\sigma_l = 2^l$ is the width of the box $B_{l,i}$.

The scale listed here is designed carefully following the parabolic-scaling principle. This is the key to the success of our multi-scale imaging process as it provides the theoretical justification of the size of each Gaussian wavepacket. The later proof and error analysis will rely on this conclusion heavily.

To have a partition of unity, one also needs the conjugate filters $h_{l,i}$, such that

$$h_{l,i}(\xi) = \frac{g_{l,i}(\xi)}{\sum_{l,i} g_{l,i}^2(\xi)}, \tag{2.17}$$

The proof that the functions $h_{l,i}$ are well defined and well-localized can be found in paper [48]. It is easy to see that $\sum_{l,i} g_{l,i} h_{l,i} = 1$. By shifting the central point,

$$\hat{\phi}_{l,i,k}(\xi) = \frac{1}{L_l^{d/2}} e^{-2\pi i \frac{k\xi}{L_l}} g_{l,i}(\xi),$$
$$\hat{\psi}_{l,i,k}(\xi) = \frac{1}{L_l^{d/2}} e^{-2\pi i \frac{k\xi}{L_l}} h_{l,i}(\xi).$$

Taking the inverse Fourier transforms yields their definitions in the spatial domain:

$$\phi_{l,i,k}(x) = \frac{1}{L_l^{d/2}} \int_{\mathbb{R}^d} e^{2\pi(x-\frac{k}{L_l})\cdot\xi} g_{l,i}(\xi) d\xi \tag{2.18}$$

$$\psi_{l,i,k}(x) = \frac{1}{L_l^{d/2}} \int_{\mathbb{R}^d} e^{2\pi(x-\frac{k}{L_l})\cdot\xi} h_{l,i}(\xi) d\xi \tag{2.19}$$

The approximation expression of the wavepacket $\phi_{l,i,k}$,

$$\phi_{l,i,k}(x) \approx \left(\sqrt{\frac{\pi}{L_l}}\sigma_l\right)^d e^{2\pi i(x-\frac{k}{L_l})\xi_{l,i}} e^{-\sigma_l^2\pi^2|x-\frac{k}{L_l}|^2}. \tag{2.20}$$

We list the lemma from paper [48] without proof to show that our decomposition is correct.

Lemma 2.2.1. *For any $f \in L_2(\mathbb{R}^d)$, we have*

$$f(x) = \sum_{l,i,k} \langle \psi_{l,i,k}, f \rangle \phi_{l,i,k}(x). \tag{2.21}$$

The idea of decomposing discrete signals into wavepackets is very similar to the continuous case, therefore, we skip this part and provide the pseudo code below. The total

Algorithm 1 Discrete Gaussian Wavepacket Decomposition

1. Apply the Fast Fourier Transform(FFT) to the discrete signal f. 2. Compute $h_{l,i}(\xi)\hat{f}(\xi)$. 3. Wrap the result to the domain $[-2\sigma_l, 2\sigma_l]$. 4. Apply the Inverse Fourier Transform to obtain the coefficients $c_{l,i,k}$.

cost of this algorithm is $O(N^d \log(N))$.

2.2.1 Multiscale Gaussian Beam Method

With the Mutiscale Gaussian Wavepacket transform defined above, the initial condition (x_0, p_0, M_0, A_0) for a single beam solution will be defined corresponding to a wavepacket $\phi_{l,i,k}$,

$$
\begin{aligned}
\frac{dx}{dt} &= G_p(x(t), p(t)), \quad x_0 = \frac{k}{L_l}; \\
\frac{dp}{dt} &= -G_x(x(t), p(t)), \quad p_0 = 2\pi \frac{\xi_{l,i}}{|\xi_{l,i}|}, \\
\frac{dM}{dt} &= -G_{xx} - MG_{xp} - G_{xp}^T M - MG_{pp}M, \quad M_0 = i(2\pi^2 \sigma_l^2 / |\xi_{l,i}|)I, \\
\frac{dA}{dt} &= -\frac{A}{2G}(v^2 trace(M) - G_x \cdot G_p - G_p^T MG_p), \quad A_0 = \left(\sqrt{\frac{\pi}{L_l N}} \sigma_l\right)^d .
\end{aligned} \tag{2.22}
$$

and

$$
u(x, t) = A(x(t), t) e^{i|\xi_{l,i}|\tau(x,t)}, \tag{2.23}
$$

Then we can use the dynamic system (2.7)-(2.12) to propagate beams. The large wavenumber $|\xi_{l,i}|$ serves as the key point of asymptotic methods. However, throughout the following derivation, we will combine this constant $|\xi_{l,i}|$ into the phase for convenience.

We argue that there's no difference. Denote the beam solution using initial condition (x_0, p_0, M_0, A_0) as $(x(t), p(t), M(t), A_t)$ and the one using initial condition $(x_0, |\xi_{l,i}|p_0, |\xi_{l,i}|M_0, A_0)$ as $(\hat{x}(t), \hat{p}(t), \hat{M}(t), \hat{A}(t))$. With respect to the ray trajectory,

$$
\begin{aligned}
\frac{dx(t)}{dt} &= v(x(t)) \frac{|\xi_{l,i}| p(t)}{|\xi_{l,i}||p(t)|}, \quad x_0 = \frac{k}{L_l} \\
\frac{d(|\xi_{l,i}| p(t))}{dt} &= \nabla v(x(t)) |\xi_{l,i}||p(t)|, \quad |\xi_{l,i}| p_0 = 2\pi \xi_{l,i}
\end{aligned}
$$

Since $(x(t), |\xi_{l,i}|p(t))$ is the solution of the same Hamiltonian system equipped with the same initial condition as $(\hat{x}(t), \hat{p}(t))$, by the uniqueness of the initial value problem, $\hat{p}(t) = |\xi_{l,i}|p(t)$ and $\hat{x}(t) = x(t)$.

In terms of Hessian M, again we multiply $|\xi_{l,i}|$ on both sides of equation (2.11).

$$
\begin{aligned}
\frac{d(|\xi_{l,i}|M)}{dt} &= -\nabla\nabla v(x(t))|\xi_{l,i}||p(t)| - |\xi_{l,i}|M\nabla v\frac{|\xi_{l,i}|p(t)}{|\xi_{l,i}||p(t)|} - |\xi_{l,i}|\nabla v\left(\frac{|\xi_{l,i}|p(t)}{|\xi_{l,i}||p(t)|}\right)^T M \\
&- (|\xi_{l,i}|M)\left(\frac{v(x(t))}{|\xi_{l,i}||p(t)|}I - \frac{v(x(t))}{|\xi_{l,i}||p(t)|^3}p(t)p_t^T\right)(|\xi_{l,i}|M) \\
&= -\nabla\nabla v(\hat{x}(t))|\hat{p}(t)| - \hat{M}G_{xp}(\hat{x}(t), \hat{p}(t)) - G_{xp}^T(\hat{x}(t), \hat{p}(t))\hat{M} \\
&- \hat{M}G_{pp}(\hat{x}(t), \hat{p}(t))\hat{M}
\end{aligned}
$$

Again, by the uniqueness of the initial value problem, $\hat{M}_t = |\xi_{l,i}|M(t)$. The derivative term G_{pp} is given by

Proposition 2.2.1. *The second order derivative of the Hamiltonian about the momentum variable is*

$$
G_{pp}^{\pm}(x,p) = \pm\frac{v(x)}{|p|^3}\left(|p|^2 I_d - pp^T\right), \tag{2.24}
$$

where I_d is the identity matrix.

The proof is easy, so we omit it here.

$$
\begin{aligned}
\frac{d\hat{A}}{dt} &= -\frac{\hat{A}}{2|\xi_{l,i}|G(x(t),p(t))}(v^2|\xi_{l,i}|M(t) - |\xi_{l,i}|G_x(x(t),p(t)) \cdot G_p(x(t),p(t))) \\
&+ \frac{\hat{A}}{2|\xi_{l,i}|G(x(t),p(t))}\left(|\xi_{l,i}|G_p^T(x(t),p(t))M(t)G_p(x(t),p(t))\right).
\end{aligned} \tag{2.25}
$$

Apparently, $A(t)$ and $\hat{A}(t)$ share the same ODE and initial condition.

Chapter 3

Multiscale Gaussian-Beam Transforms for True Amplitude Prestack Migration

3.1 Modified Multiscale Gaussian Wavepacket Transform

We have so far finished introducing the Gaussian wavepacket transform in [48]. To meet our inversion algorithm's requirements, we have to substitute $\langle \psi_{l,i,k}, f \rangle$ in equation (2.21) with $\langle \phi_{l,i,k}, f \rangle$.

Several new notations are needed in the following arguments. We first define a set for each frequency ξ as its cover set $\mathcal{S}(\xi)$,

$$\mathcal{S}(\xi) = \{(l,i) : g_{l,i}(\xi) > 0\}. \tag{3.1}$$

B_l is defined as the border of the partitioning,

$$B_l \equiv$$
$$\{\xi : \max_{1\le s\le d} |\xi_s| \in [4^{l-1}, 4^l]\} \cap \{\xi : \max_{1\le s\le d} |\xi_s| - 4^{l-1} \ge 2^l\} \cap \{\xi : 4^l - \max_{1\le s\le d} |\xi_s| \le 2^l\}. \tag{3.2}$$

The remainder part $C_l \setminus B_l$ is defined as the major part of C_l. Obviously, the border part B_l is much smaller compared with the major part. We will focus on the major part first.

All frequencies in the major part of C_l will not interact with those from different levels, that is, the cover sets of all frequency ξ in the major part $C_l \setminus B_l$ will only contain the compact support functions $g_{l,i}$ from the same level. This allows us to prove the following claims.

Proposition 3.1.1. *If $\xi \in C_l \setminus B_l$, we have*

$$\sum_{l',i'} g^2_{l',i'}(\xi) = \sum_{l'<l,i'} g^2_{l',i'}(\xi) + \sum_{l'=l,i'} g^2_{l',i'}(\xi) + \sum_{l'>l,i'} g^2_{l',i'}(\xi) = \sum_{l'=l,i'} g^2_{l',i'}(\xi). \tag{3.3}$$

Proof. We assume that the concerned frequency $\xi = (\xi_1, \xi_2, \cdots, \xi_d)$ satisfies,

$$\xi_1 = \max_{1\le s\le d} |\xi_s|. \tag{3.4}$$

We first check the first term in equation (3.3), where $l^{'} < l$,

$$\sum_{l^{'}<l,i^{'}} g^2_{l^{'},i^{'}}(\xi) = 0. \tag{3.5}$$

The nearest central frequency $\xi_{l^{'},i^{'}}$ in the lower level $l^{'} < l$ should have

$$||\xi - \xi_{l^{'},i^{'}}|| \geq |\xi_1 - \xi_{l^{'},i^{'},1}| \geq 2^l, \tag{3.6}$$

On the other hand, $g_{l^{'},i^{'}}$ is a compact support function in the box centered at $\xi_{l^{'},i^{'}}$ with side length $2^{l^{'}+1}$. Since

$$2^{l^{'}+1} \leq 2^l, \tag{3.7}$$

the frequency ξ in the level l is at most on the edge of the box $B_{l^{'}i^{'}}$.

We then check the third term in equation (3.3) where $l^{'} > l$. We denote,

$$\xi_{l^{'},i^{'},s_0} = \max_{1\leq s\leq d} |\xi_{l^{'},i^{'},s}|. \tag{3.8}$$

Therefore,

$$||\xi - \xi_{l^{'},i^{'}}|| \geq |\xi_{s_0} - \xi_{l^{'},i^{'},s_0}| \geq |\xi_{l^{'},i^{'},s_0}| - |\xi_{s_0}| \geq |\xi_{l^{'},i^{'},s_0}| - |\xi_1| \geq 2^{l+1}. \tag{3.9}$$

To summarize, for any $\xi_{l^{'},i^{'}}$ on the higher level, we have

$$|\xi_{s_0} - \xi_{l^{'},i^{'},s_0}| \geq 2^{l+1}, \tag{3.10}$$

along the s_0 direction. We notice that the compact support area of $g_{l',i'}$ along the direction s_0 has side length 2^{l+2}, however, ξ must be on the left side of $\xi_{l',i'}$ along the s_0 direction. Then, ξ will be at most on the edge of support area of the function $g_{l',i'}$. □

For any central frequency $\xi_{l,i} \in C_l \setminus B_l$, we define $J_{l,i}$

$$J_{l,i} \equiv \sum_{l',i'} g^2_{l',i'}(\xi_{l,i}). \tag{3.11}$$

By Proposition 3.1.1,

$$J_{l,i} \equiv \sum_{l'=l,i'} g^2_{l',i'}(\xi_{l,i}). \tag{3.12}$$

Proposition 3.1.2. $J_{l,i}$ *is independent of the index* i*, that is*

$$J_{l,i} = J_l = \sum_{l'=l,i'} g^2_{l',i'}(\xi_{l,i}), \quad \forall \xi_{l,i} \in C_l \setminus B_l. \tag{3.13}$$

Proof. Suppose there are two different central frequencies $\xi_{l,i}$ and ξ_{l,i^*} in the same major part of C_l. By Proposition 3.1.1, we should only consider $g_{l,i}$ from the same level.

On the other hand, in the fixed level l, the compact support area of each box function $g_{l,i}$ has side length 2^{l+1} along each dimension. In fact, each (l,i) will have overlapping support with only two other i-indexes in each direction, since the central frequency $\xi_{l,i}$ in the level l is chosen as $2^l i$ with all possible integer multi index i.

Using Proposition 3.1.1, we have

$$\sum_{l',i'} g^2_{l',i'}(\xi_{l,i}) \approx \left(\sum_{i'_1=i_1-1}^{i_1+1} \exp\left(-2\left(\frac{|\xi_{l,i,1}-\xi_{l,i',1}|}{\sigma_l}\right)^2\right)\right)^d$$
$$= \left(\sum_{i'=-1}^{i'=1} \exp\left(-2\left(\frac{i'}{\sigma_l}\right)^2\right)\right)^d$$
$$= \left(\sum_{i'_1=i^*_1-1}^{i^*_1+1} \exp\left(-2\left(\frac{|\xi_{l,i^*,1}-\xi_{l,i',1}|}{\sigma_l}\right)^2\right)\right)^d \approx \sum_{l',i'} g^2_{l',i'}(\xi_{l,i^*}). \tag{3.14}$$

Two summations are the same since Gaussian functions are only about the distance between two frequencies. □

Proposition 3.1.3. *For all frequency $\xi \in C_l \setminus B_l$,*

$$\sum_{i'\in\mathbb{Z}^d} e^{-2\left(\frac{||\xi-2^l i'||}{\sigma_l}\right)^2} =$$
$$\sum_{||2^l i'||_\infty \le 4^{l-1}} e^{-2\left(\frac{||\xi-2^l i'||}{\sigma_l}\right)^2} + \sum_{||2^l i'||_\infty > 4^l} e^{-2\left(\frac{||\xi-2^l i'||}{\sigma_l}\right)^2}$$
$$+ \sum_{4^{l-1}<||2^l i'||_\infty \le 4^l} e^{-2\left(\frac{||\xi-2^l i'||}{\sigma_l}\right)^2}$$
$$= \sum_{i'} e^{-2\left(\frac{||\xi-\xi_{l,i'}||}{\sigma_l}\right)^2} + \epsilon, \tag{3.15}$$

where $||x||_\infty \equiv \max_{1\le s\le d} |x_s|$ and $\xi_{l,i'}$ is the central frequency defined in the wavepacket transform at the level l. ϵ is a small number which can be ignored.

Proof.

$$\sum_{||2^l i'||_\infty \leq 4^{l-1}} e^{-2\left(\frac{||\xi - 2^l i'||}{\sigma_l}\right)^2} \leq$$

$$2\sigma_l^{-d} \int \cdots \int_0^{4^{l-1}-2^{l-1}} \prod_{s=1}^{s=d} e^{-\left(\frac{|\xi_s - c_s|}{\sigma_l}\right)^2} dc_1 dc_2 \cdots dc_d$$

$$\leq \sqrt{\pi}\mathrm{erfc}(3), \tag{3.16}$$

where erfc$(\cdot)$ is the complementary error function. Here we assume c is the central frequency at the lower level, therefore,

$$c_1 \leq 4^{l-1} - 2^{l-1}. \tag{3.17}$$

Meanwhile, we assume that $\xi_1 = \max_{1 \leq s \leq d} |\xi_s| \geq 4^{l-1} + 2^l$, then

$$\xi_1 - c_1 \geq 2^l + 2^{l-1}. \tag{3.18}$$

With $\sigma_l = 2^{l-1}$, we have the upper bound specified in equation (3.16).

Similarly, we have

$$\sigma_l^{-d} \sum_{l' > l, i'} e^{-\left(\frac{||\xi - \xi_{l',i'}||}{\sigma_l}\right)^2} \leq \sqrt{\pi}\mathrm{erfc}(5). \tag{3.19}$$

Then

$$\sum_{i' \in \mathbb{Z}^d} e^{-\left(\frac{||\xi - 2^l i'||}{\sigma_l}\right)^2} = \sum_{l'=l,i'} e^{-\left(\frac{||\xi - \xi_{l',i'}||}{\sigma_l}\right)^2} + \sqrt{\pi}\mathrm{erfc}(3) + \sqrt{\pi}\mathrm{erfc}(5). \tag{3.20}$$

□

Proposition 3.1.4. *For any frequency ξ in the major part $C_l \setminus B_l$, we have*

$$\sum_{l',i'} g^2_{l',i'}(\xi) \approx J_l = \sum_{l'=l,i'} g^2_{l',i'}(\xi_{l,i}), \quad \forall \xi \in C_l \setminus B_l. \tag{3.21}$$

Proof. We have already proved in Proposition 3.1.1 that any frequency ξ in the major part $C_l \setminus B_l$ satisfies,

$$\sum_{l',i'} g^2_{l',i'}(\xi) = \sum_{l'=l,i'} g^2_{l',i'}(\xi). \tag{3.22}$$

Meanwhile, according to the Poisson summation formula,

$$\left(\frac{a}{\sqrt{2\pi}}\right)^d \sum_{m \in \mathbb{Z}^d} e^{-\frac{\sigma^2}{2}||x+ma||^2} = \sum_{m \in \mathbb{Z}^d} e^{-\frac{1}{2\sigma^2}||2\pi\frac{m}{a}||^2} = const. \tag{3.23}$$

Then we have,

$$\sum_{i' \in \mathbb{Z}^d} e^{-2\left(\frac{||\xi - 2^l i'||}{\sigma_l}\right)^2} = \sum_{i' \in \mathbb{Z}^d} e^{-2\left(\frac{||\xi_{l,i} - 2^l i'||}{\sigma_l}\right)^2}. \tag{3.24}$$

Using equation (3.15),

$$\sum_{l',i'} g^2_{l',i'}(\xi) \approx \sum_{l'=l,i'} e^{-2\left(\frac{||\xi-\xi_{l',i'}||}{\sigma_{l'}}\right)^2} \approx \sum_{i'\in\mathbb{Z}^d} e^{-2\left(\frac{||\xi-2^l i'||}{\sigma_l}\right)^2}$$
$$= \sum_{i'\in\mathbb{Z}^d} e^{-2\left(\frac{||\xi_{l,i}-2^l i'||}{\sigma_l}\right)^2} \approx \sum_{l'=l,i'} e^{-2\left(\frac{||\xi_{l,i}-\xi_{l',i'}||}{\sigma_l}\right)^2} \approx J_l. \tag{3.25}$$

The proposition is proved. □

For the frequency ξ on the border of the partitioning, their sums satisfy

$$J_{l-1} \leq \sum_{l,i} g^2_{l,i}(\xi) \leq J_l. \tag{3.26}$$

Using J_l as an approximation will yield an overestimation, however, it is negligible since B_l is much smaller compared with the major area away from the border.

To summarize,

$$J_l \approx \sum_{l,i} g^2_{l,i}(\xi), \quad \forall \xi \in C_l. \tag{3.27}$$

Then,

$$h_{l,i}(\xi) \approx \frac{1}{J_l} g_{l,i}(\xi). \tag{3.28}$$

We therefore have a modified inverse wavepacket transform as the following,

Lemma 3.1.1. *For any* $f \in L_2(\mathbb{R}^d)$, *we have*

$$f(x) \approx \sum_{l,i,k} \frac{1}{J_l} \langle \phi_{l,i,k}, f \rangle \phi_{l,i,k}(x). \tag{3.29}$$

To end this part, we would like to display some numerical results to justify Lemma 3.1.1. The denominator in the expression of $h_{l,i}$ (2.17) should be a step function about the index l, suggested by Lemma 3.1.1. The range of the frequency concerned is $[64 : 320] \times [64 : 320]$.

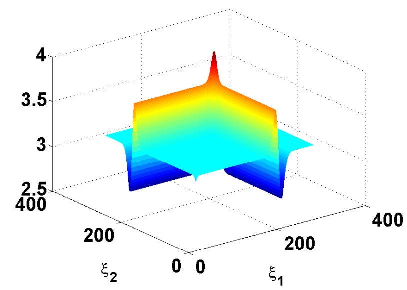


Figure 3.1: $\sum_{l,i} g_{l,i}^2$ in different level l

3.2 Multiscale Gaussian Wavepacket Inversion

3.2.1 Setup of the True Amplitude Migration Problem

Let us suppose the wave propagation is governed by the scalar wave equation (2.1) with the wave propagation velocity decomposed as,

$$c = v(1 + \alpha), \tag{3.30}$$

where v is the macro velocity being responsible for the traveltime and amplitude. Moreover, it is assumed to be smooth and does not provide the significant energy back to the boundary data. The rapid perturbation α is small but reflects the wave signal back to the boundary data. In our inversion model, v is known at prior and our target is to image α.

We simplify our model as a constant density fluid occupying a half-space $X \equiv \{x \in \mathbb{R}^d : x_d \geq 0\}$. The boundary data $D(r,s,t)$ used in this paper is organized as the common-shot trace, for example, Fig. 3.2, r parametrizes receiver positions on the surface $\partial X \equiv \{x \in \mathbb{R}^d : x_d = 0\}$, while s parametrizes source positions. We also assume that the sources are contained in an open set $O_s \subset \partial X$ and receivers are in an open set $O_r \subset \partial X$. Therefore, the boundary data D is a function defined at $O_s \times O_r \times [0,T]$. Similar to [44], we make some assumptions about rays.

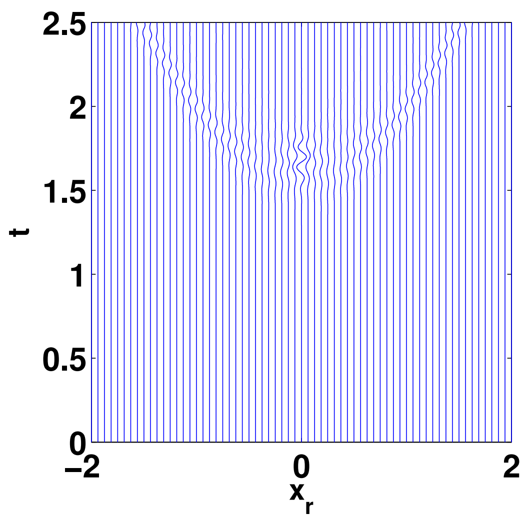


Figure 3.2: A typical source gather in Gaussian slowness

Assumption 3.2.1. *There are no rays leaving points in the subsurface $\{x \in X : x_d > 0\}$ and returning to graze O_s or O_r. Moreover, there exists an universal lower bound $\mathfrak{b}$, such that*

$$|p_d| \geq \mathfrak{b}||p||, \tag{3.31}$$

for any rays hitting the surface where p is the momentum variable and p_d is the d^{th} component of p.

Assumption 3.2.2. *Rays departing from a source in O_s and traveling into the subsurface do not return to receivers in O_r.*

Assumption 3.2.3. *There exists $\delta > 0$, such that $v(x)$ is a constant if $0 \leq x_d \leq \delta$.*

3.2.2 Born Approximation for the Trace Data

Denote the wave propagator with background velocity v as L_0 and the wave propagator with true velocity c as L. Meanwhile, the corresponding Green's functions are written as G_0 and G, respectively,

$$G = -L^{-1}; \quad G_0 = -L_0^{-1}, \tag{3.32}$$

and by some formal computations,

$$G = G_0 + G_0(L - L_0)G = G_0 + G_0 V G, \tag{3.33}$$

where $V = -L_0 + L$. The Born approximation assumes the whole scattering process as the following. The signal is initiated from the source and travel through a smooth medium afterwards. Then at some moment, it hits the reflector under the surface and is reflected back. Therefore, the boundary data $D(r, s, t)$ is the reflection data, or the wavefield along the boundary minus the direct wave. During the whole process, we assume that the reflection or scattering only happens once so that we ignore multiple reflections.

Define the incident wavefield $G_0(x, t; s)$ generated by the source point s,

$$L_0 G_0(x, t; s) = -\delta(t)\delta(x - s), \quad G_0|_{t<0} = 0. \tag{3.34}$$

Then the reflection signal is obtained by total differentiation. We assume the true velocity $c = v + \alpha v = v + \delta v$ and the total wavefield $u = u_0 + \delta u$, where $L_0 u_0 = 0$,

$$\frac{1}{c^2}\partial_t^2 u - \Delta u \approx (\frac{1}{v^2} - \frac{2\delta v}{v^3})\partial_t^2 (u_0 + \delta u) - \Delta(u_0 + \delta u). \tag{3.35}$$

Considering the first order term, we have

$$L_0 \delta u = \frac{2\delta v}{v^3}\partial_t^2 u_0. \tag{3.36}$$

Now the perturbed wavefield δG [7, 46] satisfies,

$$L_0 \delta G = \frac{2\alpha}{v^2}\frac{\partial^2 G_0}{\partial t^2}. \tag{3.37}$$

This perturbed wavefield δG is the data recorded along the surface based on the Born approximation, that is,

$$D\left[\frac{2\alpha}{v^2}\right](r, s, t) = \delta G = \int dx \frac{2\alpha}{v^2} \int dh \hat{G}_0(r, t - h; x)\frac{\partial^2 \tilde{G}_0}{\partial t^2}(x, h; s), \tag{3.38}$$

where both $\tilde{G}$ and $\hat{G}$ are Green's functions. The Green's function $\hat{G}_0(r, t - h; x)$ represents the perturbation received at the receiver r at the moment $t - h$, and its source is the subsurface point x. $\tilde{G}$ is about source points s and $\hat{G}$ is about receivers

r. The same rules are applied to other functions.

$$D\left[\frac{2\alpha}{v^2}\right](r,s,t) = \frac{\partial^2}{\partial t^2}\int dx \frac{2\alpha}{v^2}\int dh \hat{G}_0(r,t-h;x)\tilde{G}_0(x,h;s). \tag{3.39}$$

To make things easier, we would like to develop our algorithm in the frequency domain instead of the time domain so that we can simplify the convolution operator above. Applying the following Fourier transform in time, we have,

$$D\left[\frac{2\alpha}{v^2}\right](r,s,\omega) = -\omega^2\int dx \frac{2\alpha}{v^2}\hat{G}_0(r,\omega;x)\tilde{G}_0(s,\omega;x). \tag{3.40}$$

Here we abuse the notation by writing the Fourier transform of the Green's function $\hat{G}_0(r,t;x)$ about the time variable t by $\hat{G}_0(r,\omega;x)$. In equation (3.40), the reciprocity of the Green's function is involved as we replace $\tilde{G}_0(x,\omega;s)$ with $\tilde{G}_0(s,\omega;x)$.

3.2.3 Multiscale Gaussian Beam Approximation of the Green's Function

We then approximate the Green's function in the high frequency regime by the summation of Gaussian beams so that we can define the following multiscale Gaussian-beam transform of the perturbation of the velocity,

$$D\left[\frac{2\alpha}{v^2}\right](r,s,\omega) = -\omega^2\int dx \frac{2\alpha}{v^2}\iint d\xi d\eta \hat{U}_{GB}(r,\omega;x,\xi)\tilde{U}_{GB}(s,\omega;x,\eta), \tag{3.41}$$

where without confusion, we sometimes shorten $D\left[\frac{2\alpha}{v^2}\right](r,s,\omega)$ to be $D(r,s,\omega)$ so that we use $D(r,s,\omega)$ to denote the Fourier transform of the boundary data D with respect to the time variable t. $\hat{U}_{GB}(r,\omega;x,\xi)$ is the beam solution in the frequency domain starting at the point x with the initial momentum ξ. From now on, the following notation is used to the end,

$$\hat{U}_{GB}(r,\omega;x,\xi) = \hat{U}^x_\xi(r,\omega); \quad \tilde{U}_{GB}(s,\omega;x,\eta) = \tilde{U}^x_\eta(s,\omega).$$

The Green's function can be considered as the solution to the acoustic wave equation whose initial velocity is a Dirac-delta function by Duhamel's principle and the multiscale Gaussian wavepacket transform can be applied to decompose the Dirac-delta function.

$$\frac{1}{v^2(x)}\partial_t^2 G_0(x,t;s) - \Delta G_0(x,t;s) = 0, \quad G_0\Big|_{t=0} = 0, \quad \frac{\partial G_0}{\partial t}\Big|_{t=0} = -\delta(t)\delta(x-s). \tag{3.42}$$

Although the multiscale transform introduced in Section 2.2 and reference [48] is designed to decompose any general L_2 functions, the Dirac-delta function can be approximated by some L_2 functions.

3.2.4 True-Amplitude Migration Process

A new operator K_{pq} applied to the perturbation α can be defined, which is corresponding to the certain pair of Gaussian beams,

$$\left(K_{pq}\frac{2\alpha}{v^2}\right)(r,s,\omega) = \int dx \frac{2\alpha}{v^2}\hat{U}^x_p(r,\omega)\tilde{U}^x_q(s,\omega), \tag{3.43}$$

which will be called the atomic Gaussian-beam transform. The operator K_{pq} maps the subsurface information to the boundary data.

$$(K^*_{pq}g)(y,\omega) = \iint drds \bar{\hat{U}}^y_p \bar{\tilde{U}}^y_q g(r,s,\omega), \tag{3.44}$$

which is the adjoint of the atomic Gaussian-beam transform. Applying K^*_{pq} to the boundary data D yields a single-frequency prestack angle-gather imaging function I_{pq},

$$I_{pq}(y,\omega) = \left(K^*_{pq}D\right)(y,\omega), \tag{3.45}$$

or to write it completely,

$$I_{pq}\left[\frac{2\alpha}{v^2}\right](y,\omega) = \left(K^*_{pq}D\left[\frac{2\alpha}{v^2}\right]\right)(y,\omega).$$

Substitute equation (3.41) about the surface data D into equation (3.45),

$$\begin{aligned} I_{pq}\left[\frac{2\alpha}{v^2}\right](y,\omega) = -\omega^2 \iint drds \bar{\hat{U}}^y_p(r,\omega)\bar{\tilde{U}}^y_q(s,\omega) \\ \int dx \frac{2\alpha}{v^2} \iint d\xi d\eta \hat{U}^x_\xi(r,\omega)\tilde{U}^x_\eta(s,\omega). \end{aligned} \tag{3.46}$$

We will later show that the following approximation (3.47) is correct,

$$\begin{aligned} &-\iiiint d\omega drdsd\xi d\eta \omega^2 \bar{\hat{U}}^y_p(r,\omega)\bar{\tilde{U}}^y_q(s,\omega)\hat{U}^x_\xi(r,\omega)\tilde{U}^x_\eta(s,\omega) \\ &\approx E(p,q,y)e^{i(p+q)(y-x)}e^{(y-x)^T\frac{iM_0}{2}(y-x)}, \end{aligned} \tag{3.47}$$

where $E(p,q,y)$ is a constant related to the parameters of the corresponding wavepacket, and M_0 is a symmetric matrix with a positive definite imaginary part as defined in the Multiscale Gaussian beam propagation. The integral $\int d\omega$ should be interpreted as $\int \chi(\omega)d\omega$, where χ is an arbitrary C^∞ function which is zero in the low frequency region and is 1 for the high frequency.

If we integrate (3.46) further with respect to ω, then we will have the wavepacket transform about $\frac{2\alpha}{v^2}$,

$$\int I_{pq}\left[\frac{2\alpha}{v^2}\right](y,\omega)d\omega = E(p,q,y)\int dx\frac{2\alpha}{v^2}e^{i(p+q)(y-x)}e^{i(y-x)^T\frac{M_0}{2}(y-x)}. \tag{3.48}$$

As we can see in equation (3.48), $\int d\omega I_{pq}$ is essentially the Gaussian wavepacket transform of the function $\frac{2\alpha}{v^2}$ in the direction of $p+q$. The integral $\int d\omega I_{pq}(y,\omega)$ can be considered as taking the inverse Fourier transform to obtain $I_{pq}(y,t)$ at $t=0$. According to Claerbout imaging principle [29], $I_{pq}(y,t)$ at $t=0$ yields the initial state of the subsurface that we want to image.

By using the modified wavepacket transform (3.29), we can reconstruct perturbation $\frac{2\alpha}{v^2}$ through the imaging function (3.48).

$$\begin{aligned}
\frac{2\alpha}{v^2} &= \sum_{l,i,k}\frac{1}{J_l}\langle\frac{2\alpha}{v^2},\phi_{l,i,k}\rangle\phi_{l,i,k}(x) \\
&= \sum_{l,i,k}\sum_{p+q=\xi_{l,i}}\frac{\int I_{pq}\left[\frac{2\alpha}{v^2}\right](\frac{k}{L_l},\omega)d\omega}{EJ_l}\phi_{l,i,k}(x).
\end{aligned} \tag{3.49}$$

Based on equation (3.49) ,

Algorithm 2 Multiscale Gaussian Beam True-Amplitude Migration

1.Run the multiscale Gaussian wavepacket transform to get the dictionary of central points y and central momentum $p+q$.

2. Separate p+q into two different wavepackets and shoot them from the subsurface point y to the acquisition surface.

3. Angle-gather prestack image by calculating $I_{pq}(y,\omega) = \iint dsdr \bar{\hat{U}}_p^y \bar{\tilde{U}}_q^y D(r,s,\omega)$.

4. Stack all partial image I_{pq} with the same $p+q$ to reduce the noise.

5. Use equation (3.48) to get the coefficient of each wavepacket expansion of the perturbation $\frac{2\alpha}{v^2}$.

6. Run the inverse multiscale Gaussian wavepacket transform to recover the rapid perturbation $\frac{2\alpha}{v^2}$.

3.2.5 Motivation for Inverting the multiscale Gaussian-beam transform

Hereby we provide some intuitive justifications of the inversion process of the multiscale Gaussian-beam transform, which may provide some theoretical guideline for carrying out further analysis of our new methodology and extending our methodology to other applications.

We start with considering the function $b(r,s,t)$ defined by the linear operator D,

$$D[f](r,s,t) = b(r,s,t), \tag{3.50}$$

where the function f is defined at subsurface points. To solve this linear operator equation rapidly and efficiently, we would like to diagonalize the operator D by carrying out a certain frame representation. Since the argument f of the operator D does not sit in the same space as the right-hand side b, we first apply the adjoint operator D^* to both sides, so that we have

$$D^* D[f] = D^* b. \tag{3.51}$$

Let $\mathcal{W}$ be the multiscale Gaussian wavepacket transform defined in Section 2.2, which satisfies $\mathcal{W}^*\mathcal{W} = I$. Then we apply $\mathcal{W}$ to both sides of equation (3.51), yielding

$$\mathcal{W}(D^* D)\mathcal{W}^*\mathcal{W} f = \mathcal{W} D^* b. \tag{3.52}$$

Our results in Section 3.2.4 indicate that the above diagonalization is justified. The operator $\mathcal{W} D^* b$ is essentially equation (3.46), and the operator $\mathcal{W}(D^* D)\mathcal{W}^*$ is essentially captured by the diagonal factor $E(p, q, y)$ in equation (3.47). The overall effects of the diagonalization process are epitomized in equation (3.48).

Therefore, we have

$$\begin{aligned} \mathcal{W} f &= \left(\mathcal{W}(D^* D)\mathcal{W}^*\right)^{-1} \mathcal{W} D^* b, \\ f &= \mathcal{W}^* \left(\mathcal{W}(D^* D)\mathcal{W}^*\right)^{-1} \mathcal{W} D^* b, \end{aligned} \tag{3.53}$$

which results in our fast reconstruction formula (3.49).

To establish the theoretical foundation for our new methodology in terms of the Fourier Integral Operator (FIO) theory, we need to carry out symbolic calculus to establish several facts by following the works in [50, 23]:

1. the forward operator D belonging to a certain class of FIO;

2. the conjugation process $\mathcal{W}(D^*D)\mathcal{W}$ yielding a diagonal operator in the frame defined by $\mathcal{W}$.

On the other hand, since the FIO theory is originated from the asymptotic of geometric optics and a Gaussian-beam solution provides a globally defined asymptotic solution for wave equations, we will carry out brute-force calculations to justify our new methodology by heavily relying on the structure of multiscale Gaussian-beam transform, which is composed of Gaussian beams and multiscale Gaussian wavepacket transform.

3.3 Theoretical Validation: The Proof about The Imaging Operator

3.3.1 Road Map of The Theoretical Analysis

We will prove the approximation (3.47) holds. Since it involves the interaction of four beams in time, we will carry out the analysis essentially in two main steps.

The first step consists of analysis of the following two integrals dealing with the beam interactions on the receiver side and the source side respectively,

$$\iint drd\xi \bar{\hat{U}}_p^y(r,\omega)\hat{U}_\xi^x(r,\omega), \quad \iint dsd\eta \bar{\tilde{U}}_q^y(s,\omega)\tilde{U}_\eta^x(s,\omega).$$

Since receivers and sources are reciprocal in wave propagation, we just need to focus on analysis of beam interaction on the receiver side, and the analysis of the source side is essentially analogous.

Our analysis of the two beams' interaction yields the following approximations

$$\iint dr d\xi \bar{\hat{U}}_p^y(r,\omega)\hat{U}_\xi^x(r,\omega) \approx e^{ip\cdot(y-x)}\hat{H}(x,\xi,\omega;y,p),$$
$$\iint ds d\eta \bar{\tilde{U}}_q^y(s,\omega)\tilde{U}_\eta^x(s,\omega) \approx e^{iq\cdot(y-x)}\tilde{H}(x,\eta,\omega;y,q).$$

where the functions $\hat{H}$ and $\tilde{H}$ will be defined later.

Based on the first step, the second step consists of analyzing the left-hand side of approximation (3.47) so that we have

$$\iiiint dr ds d\xi d\eta \bar{\hat{U}}_p^y(r,\omega)\bar{\tilde{U}}_q^y(s,\omega)\hat{U}_\xi^x(r,\omega)\tilde{U}_\eta^x(s,\omega)$$
$$\approx e^{i(p+q)(y-x)}\hat{H}(x,\xi,\omega;y,p)\tilde{H}(x,\eta,\omega;y,q). \tag{3.54}$$

After carrying out the integral about ω, the above approximation reduces to

$$-\iiiint d\omega dr ds d\xi d\eta \omega^2 \bar{\hat{U}}_p^y(r,\omega)\bar{\tilde{U}}_q^y(s,\omega)\hat{U}_\xi^x(r,\omega)\tilde{U}_\eta^x(s,\omega) =$$
$$e^{i(p+q)(y-x)}K(p,q,y)\hat{\mathcal{H}}(x,y,p,q)\tilde{\mathcal{H}}(x,y,p,q). \tag{3.55}$$

where $K(p,q,y)$, $\hat{\mathcal{H}}$ and $\tilde{\mathcal{H}}$ are defined later.

After making an essential assumption on the invertibility of the imaging operator, we are able to approximate functions $\hat{\mathcal{H}}$ and $\tilde{\mathcal{H}}$ by

$$\hat{\mathcal{H}}(x,y,p,q) \approx e^{\frac{i}{4}||y-x||^2_{\hat{M}(0)}} \hat{\mathcal{K}}(y,p,q),$$
$$\tilde{\mathcal{H}}(x,y,p,q) \approx e^{\frac{i}{4}||y-x||^2_{\tilde{M}(0)}} \tilde{\mathcal{K}}(y,p,q),$$

where $\hat{\mathcal{K}}(y,p,q)$ and $\tilde{\mathcal{K}}(y,p,q)$ can be easily computed.

These latter approximations allow us to obtain our main theorem,

$$\int -\omega^2 d\omega \int drds \bar{\hat{U}}^y_p(r,\omega)\bar{\tilde{U}}^y_q(s,\omega) \int d\xi d\eta \hat{U}^x_\xi(r,\omega)\tilde{U}^x_\eta(s,\omega) \approx$$
$$K(p,q,y)\hat{\mathcal{K}}(y,p,q)\tilde{\mathcal{K}}(y,p,q)e^{i(p+q)\cdot(y-x)}e^{i||y-x||^2_{M_0/2}},$$

which says that the four-beam interaction in time yields a weighted Gaussian wavepacket centered at the scattering point y in the direction $p+q$, where (y,p) is the ray parameter for the beam from the scattering point y to the boundary receiver in the direction p, and (y,q) is the ray parameter for the beam from the scattering point y to the boundary source in the direction q.

The rest of the analysis will follow the above road map.

We will prove equation (3.47) in this section. Throughout the proof, the amplitude function A is not involved as it is more smooth compared with the phase function part. Denote the beam as $(\hat{y}(t), \hat{p}(t), \hat{M}(t), \hat{A}(t))$, whose initial value is (y,p,M,A), and $(\hat{x}(t), \hat{\xi}(t), \hat{N}(t), \hat{C}(t))$ whose initial value is (x,ξ,N,C). Moreover, $\hat{\Xi}$ and $\hat{\kappa}$ are defined according to a fixed beam $\hat{U}_p^y$,

$$\hat{\xi}(t) = \hat{\kappa}(t; x, \xi, y, p)\hat{p}(t) + \hat{\Xi}(t; x, \xi, y, p), \tag{3.56}$$

and

$$\hat{\Xi}(t; x, \xi, y, p) \cdot \hat{p}(t) = 0. \tag{3.57}$$

On the source side, we have similar notation $(\tilde{y}(t), \tilde{q}(t), \tilde{M}(t), \tilde{A}(t))$, whose initial value is (y, q, M, A), and $(\tilde{x}(t), \tilde{\eta}(t), \tilde{N}(t), \tilde{C}(t))$ whose initial value is (x, η, N, C).

3.3.2 Approximation of Gaussian Beams along the Surface

The beam functions $\hat{U}_p^y(r, \omega)$ and $\tilde{U}_p^y(s, \omega)$ are used in the inversion process to link the data to the unknown perturbation. We would like to explore more about the beam function's structure to build the foundation for the future proof and calculation.

$||x||_M^2$ denotes $x^T M x$ in this paper for all vectors x and all symmetric matrices M.

For each beam $\hat{U}_p^y$, we define the hitting time $\hat{t}_0 = \hat{t}_0(y, p)$ and hitting point $\hat{r}_0 = \hat{y}(\hat{t}_0(y, p))$ according to the arrival time of its central ray at the boundary so that $\hat{r}_0 = \hat{y}(\hat{t}_0(y, p)) \in \{x \in \mathbb{R}^d : x_d = 0\}$.

In [49], the single beam is propagating by treating the time variable same as the spatial components, see also [24]. The complete Hessian matrix is S.P.D. along the directions transversal to the ray direction. Therefore, by Assumption 3.2.1, we have a Gaussian profile if intersecting the beam solution at the surface $x_d = 0$.

Proposition 3.3.1. *The Gaussian beam $\hat{U}_p^y(r, t)$ along the surface for $r \in \{x =$*

$(x_{1,\cdots,d}) : x_d = 0\}$ *can be represented as*

$$\hat{U}_p^y(r,t) \approx \hat{A}(\hat{t}_0) \exp\left(i\left(\hat{\tau}_x(\hat{t}_0;y,p)\cdot(r-\hat{r}_0) + \hat{\tau}_t(\hat{t}_0;y,p)(t-\hat{t}_0)\right)\right) \times$$
$$\exp\left(i\left(||r-\hat{r}_0||^2_{\frac{\hat{M}(\hat{t}_0)}{2}} + \frac{1}{2}\hat{\tau}_{tt}(\hat{t}_0;y,p)(t-\hat{t}_0)^2 + (t-\hat{t}_0)\hat{\tau}_{tx}^T(\hat{t}_0;y,p)(r-\hat{r}_0)\right)\right). \tag{3.58}$$

where $\hat{t}_0 = \hat{t}_0(y,p)$ *and all partial derivatives about the phase function* τ *are on the central ray. The proposition is equivalent to intersecting the complete* (t,x) *beam ansatz at the surface* $x_d = 0$.

We first introduce the way to compute the terms $\hat{\tau}_{tx}$ and $\hat{\tau}_{tt}$, which can be obtained by inserting corresponding ray parameters into equations (4.13)-(4.14).

$$\hat{\tau}_t(t;y,p) = -\hat{G}(\hat{y}(t),\hat{p}(t)), \tag{3.59}$$

$$\hat{\tau}_{tx}(t;y,p) = -\hat{G}_x(\hat{y}(t),\hat{p}(t)) - \hat{M}(t;y,p)\hat{G}_p(\hat{y}(t),\hat{p}(t)), \tag{3.60}$$

$$\hat{\tau}_{tt}(t;y,p) = -\hat{G}_p(\hat{y}(t),\hat{p}(t))\cdot\hat{\tau}_{tx}(t;y,p), \tag{3.61}$$

where $(\hat{y}(t),\hat{p}(t)) = (\hat{y}(t;y,p),\hat{p}(t;y,p))$. We also denote $\hat{M}^*$ as the complete Hessian matrix about (t, x) at $(t; y, p)$ and $t = \hat{t}_0(y, p)$,

$$
\begin{aligned}
\hat{M}^*(t; y, p) &= \begin{bmatrix} \hat{\tau}_{tt}(t; y, p), & \hat{\tau}_{tx}^T(t; y, p) \\ \hat{\tau}_{tx}(t; y, p), & \hat{M}(t; y, p) \end{bmatrix}, \\
\hat{F}(r, t; y, p) &= \hat{\tau}_t(t; y, p) + Re(\hat{\tau}_{tx}(t; y, p)) \cdot (r - \hat{y}(t)), \\
\hat{\theta}(r, t; y, p) &= \hat{p}(t) \cdot (r - \hat{y}(t)) + ||r - \hat{y}(t)||^2_{\frac{Re(\hat{M}(t))}{2}}, \\
\hat{Q}(r, t; y, p) &= -\frac{(Im(\hat{\tau}_{tx})(t; y, p))^T (r - \hat{y}(t))}{Im(\hat{\tau}_{tt})(t; y, p)}.
\end{aligned}
$$

Similarly, we denote $(\hat{x}(t), \hat{\xi}(t)) = (\hat{x}(t; x, \xi), \hat{\xi}(t; x, \xi))$, $\hat{N}^*(t; x, \xi)$, $\hat{F}(r, t; x, \xi)$, $\hat{\theta}(r, t; x, \xi)$ and $\hat{Q}(r, t; x, \xi)$.

The following proposition is needed when taking the Fourier transform of $\hat{U}_p^y(r, t)$ about time t,

Proposition 3.3.2. *For any complex number γ with positive real part, i.e. $Re(\gamma) > 0$,*

$$
\int_{-\infty}^{\infty} e^{-\gamma t^2} e^{-i\omega t} dt = \sqrt{\frac{\pi}{\gamma}} e^{-\frac{\omega^2}{4\gamma}}. \tag{3.62}
$$

Lemma 3.3.1. *The Fourier transform of $\hat{U}_p^y(r, t)$ with respect to time t is*

$$
\begin{aligned}
\hat{U}_p^y(r, \omega) &= \hat{U}_p^y(r, \omega; \mathsf{t})\Big|_{\mathsf{t}=\hat{t}_0(y,p)} \\
&= e^{i\hat{\varrho}(r,\mathsf{t};y,p)} e^{-i\omega \mathsf{t}} e^{i\hat{\beta}(\mathsf{t};y,p)|\omega - \hat{\tau}_t(\mathsf{t};y,p) - \hat{\zeta}(\mathsf{t};y,p)^T (r - \hat{y}(\mathsf{t}))|^2} \\
&\quad \sqrt{\frac{i2\pi}{\hat{\tau}_{tt}(\mathsf{t}; y, p)}} \hat{A}(\mathsf{t}) e^{-||r - \hat{y}(\mathsf{t})||^2_{\frac{\hat{\mathcal{M}}(\mathsf{t};y,p,M_0)}{2}}},
\end{aligned} \tag{3.63}
$$

where

$$\hat{\mathcal{M}}(t;y,p) = Im(\hat{M})(t;y,p) - \frac{Im(\hat{\tau}_{tx})Im(\hat{\tau}_{tx})^T}{Im(\hat{\tau}_{tt}(t;y,p))} \tag{3.64}$$

$$\hat{\beta}(t;y,p) = -\frac{1}{2\hat{\tau}_{tt}(t;y,p)} = -\frac{1}{2iIm(\hat{\tau}_{tt}(t;y,p)) + 2Re(\hat{\tau}_{tt}(t;y,p))}, \tag{3.65}$$

$$\hat{\zeta}(t;y,p) = Re(\hat{\tau}_{tx}(t;y,p)) - \frac{Re(\hat{\tau}_{tt}(t;y,p))}{Im(\hat{\tau}_{tt}(t;y,p))} Im(\hat{\tau}_{tx}(t;y,p)), \tag{3.66}$$

$$\begin{aligned}\hat{\varrho}\,(r,t;y,p) &= \hat{\theta}(r,t;y,p) + (\hat{F}(r,t;y,p) - \omega)\hat{Q}(r,t;y,p) \\ &\quad + \frac{Re(\hat{\tau}_{tt}(t;y,p))}{2}\hat{Q}(r,t;y,p)^2.\end{aligned} \tag{3.67}$$

Here $\mathfrak{t}$ *in* $\hat{U}^y_p(\cdot;\mathfrak{t})$ *serves as a fixed parameter, since all terms defined above is defined at this fixed moment.*

Proof. We still abbreviate $\hat{t}_0(y,p)$ as $\hat{t}_0$ in this proof since we only concern the single beam $\hat{U}^y_p$ here. From Proposition 3.3.1,

$$\begin{aligned}\hat{U}^y_p(r,\omega) &\approx \hat{A}(\hat{t}_0)e^{i\hat{\theta}(r,\hat{t}_0;y,p)} \int e^{-i\omega t} e^{i\hat{F}(r,\hat{t}_0;y,p)(t-\hat{t}_0)} e^{i\frac{1}{2}Re(\hat{\tau}_{tt}(\hat{t}_0;y,p))(t-\hat{t}_0)^2} \\ &\quad \times e^{-||(t-\hat{t}_0,r-\hat{y}(\hat{t}_0))||^2_{Im(\hat{M}^*(\hat{t}_0;y,p,M_0))/2}} dt,\end{aligned} \tag{3.68}$$

After expanding the term $e^{-||(t-\hat{t}_0,r-\hat{y}(\hat{t}_0))||^2_{Im(\hat{M}^*/2)}}$, we have

$$\begin{aligned}&\hat{U}^y_p(r,\omega) \approx \hat{A}(\hat{t}_0)e^{i\hat{\theta}}e^{-i\omega\hat{t}_0}e^{-||(0,r-\hat{y}(\hat{t}_0))||^2_{\frac{Im(\hat{M}^*)(\hat{t}_0;y,p,M_0)}{2}}} \\ &\int e^{-i(-\hat{F}+\omega)(t-\hat{t}_0)} e^{i\frac{1}{2}Re(\hat{\tau}_{tt}(\hat{t}_0;y,p))(t-\hat{t}_0)^2} e^{-\frac{Im(\hat{\tau}_{tt}(\hat{t}_0;y,p))}{2}(t-\hat{t}_0)^2} \\ &e^{-(t-\hat{t}_0)(Im(\hat{\tau}_{tx})(\hat{t}_0;y,p))^T(r-\hat{y}(\hat{t}_0))} dt.\end{aligned} \tag{3.69}$$

To make a complete square term in equation (3.69), we have

$$-\frac{Im(\hat{\tau}_{tt})(\hat{t}_0;y,p)}{2}(t-\hat{t}_0)^2-(t-\hat{t}_0)Im(\hat{\tau}_{tx}(\hat{t}_0;y,p))^T(r-\hat{y}(\hat{t}_0))=$$
$$\frac{|(Im(\hat{\tau}_{tx})(\hat{t}_0;y,p))^T(r-\hat{y}(\hat{t}_0))|^2}{2Im(\hat{\tau}_{tt}(\hat{t}_0;y,p))}-\frac{Im(\hat{\tau}_{tt})}{2}\left(t-\hat{t}_0+\frac{(Im(\hat{\tau}_{tx}))^T(r-\hat{y}(\hat{t}_0))}{Im(\hat{\tau}_{tt})}\right)^2.$$

Since $\hat{Q}(r,\hat{t}_0;y,p)=-\frac{(Im(\hat{\tau}_{tx})(\hat{t}_0;y,p))^T(r-\hat{y}(\hat{t}_0))}{Im(\hat{\tau}_{tt})(\hat{t}_0;y,p)}$,

$$\hat{U}_p^y(r,\omega)=\hat{A}(\hat{t}_0)e^{i\hat{\theta}}e^{-i\omega\hat{t}_0}e^{i(\hat{F}-\omega)\hat{Q}}e^{i\frac{1}{2}Re(\hat{\tau}_{tt})\hat{Q}^2}$$
$$e^{\frac{|(Im(\hat{\tau}_{tx})(\hat{t}_0;y,p))^T(r-\hat{y}(\hat{t}_0))|^2}{2Im(\hat{\tau}_{tt})(\hat{t}_0;y,p)}}e^{-||r-\hat{y}(\hat{t}_0)||^2\frac{Im(\hat{M})(\hat{t}_0)}{2}}\int e^{i(\hat{F}-\omega+Re(\hat{\tau}_{tt})\hat{Q})(t-\hat{t}_0-\hat{Q})}$$
$$e^{i\frac{1}{2}Re(\hat{\tau}_{tt})(t-\hat{t}_0-\hat{Q})^2}e^{-\frac{Im(\hat{\tau}_{tt})(\hat{t}_0;y,p)}{2}(t-\hat{t}_0-\hat{Q})^2}dt+O(|t-t_0|^2)$$
$$\approx e^{i\hat{F}\hat{Q}}e^{i\frac{1}{2}Re(\hat{\tau}_{tt})(\hat{Q})^2}e^{-i\omega(\hat{t}_0+\hat{Q})}e^{\frac{|(Im(\hat{\tau}_{tx})(\hat{t}_0;y,p))^T(r-\hat{y}(\hat{t}_0))|^2}{2Im(\hat{\tau}_{tt})(\hat{t}_0;y,p)}}e^{-||r-\hat{y}(\hat{t}_0)||^2\frac{Im(\hat{M})(\hat{t}_0)}{2}}$$
$$\times\sqrt{\frac{i2\pi}{\hat{\tau}_{tt}(\hat{t}_0;y,p)}}\hat{A}(\hat{t}_0)e^{i\hat{\theta}}e^{-\hat{\epsilon}|-\hat{\tau}_t+\omega-Re(\hat{\tau}_{tx})(\hat{t}_0;y,p)\cdot(r-\hat{y}(\hat{t}_0))-Re(\hat{\tau}_{tt})(\hat{t}_0;y,p)\hat{Q}|^2}. \quad (3.70)$$

Equation (3.70) is obtained by Proposition 3.3.2 where

$$\hat{\epsilon}(\hat{t}_0;y,p)=\frac{1}{2Im(\hat{\tau}_{tt})(\hat{t}_0;y,p)-i2Re(\hat{\tau}_{tt})(\hat{t}_0;y,p)}. \quad (3.71)$$

After replacing $\hat{\beta}=i\hat{\epsilon}$, we have the lemma proved. □

Corollary 3.3.1. *The Fourier transform of $\hat{U}^x_\xi(r,t)$ with respect to time t is*

$$
\begin{aligned}
\hat{U}^x_\xi(r,\omega) &= \hat{U}^x_\xi(r,\omega;\mathfrak{t})\Big|_{\mathfrak{t}=\hat{t}_0(x,\xi)} \\
&= e^{i\hat{\varrho}(r,\mathfrak{t};x,\xi)}e^{-i\omega\mathfrak{t}}e^{i\hat{\gamma}(\mathfrak{t};x,\xi)|\omega-\hat{\tau}_t(\mathfrak{t};x,\xi)-\hat{\vartheta}(\mathfrak{t};x,\xi)^T(r-\hat{x}(\mathfrak{t}))|^2} \\
&\sqrt{\frac{i2\pi}{\hat{\tau}_{tt}(\mathfrak{t};x,\xi)}}\hat{C}(\mathfrak{t})e^{-||r-\hat{x}(\mathfrak{t})||^2\frac{\hat{\mathcal{N}}(\mathfrak{t};x,\xi,N_0)}{2}}, \qquad (3.72)
\end{aligned}
$$

where

$$
\hat{\mathcal{N}}(t;x,\xi) = Im(\hat{N})(t;x,\xi) - \frac{Im(\hat{\tau}_{tx})Im(\hat{\tau}_{tx})^T}{Im(\hat{\tau}_{tt}(t;x,\xi))} \qquad (3.73)
$$

$$
\begin{aligned}
\hat{\varrho}\,(r,t;x,\xi) &= \hat{\theta}(r,t;x,\xi) + (\hat{F}(r,t;x,\xi)-\omega)\hat{Q}(r,t;x,\xi) \\
&+ \frac{1}{2}Re(\hat{\tau}_{tt}(t;x,\xi))\hat{Q}(r,t;x,\xi)^2, \qquad (3.74) \\
\hat{\gamma}(t;x,\xi) &= -\frac{1}{2iIm(\hat{\tau}_{tt}(t;x,\xi))+2Re(\hat{\tau}_{tt}(t;x,\xi))}, \\
\hat{\vartheta}(t;x,\xi) &= Re(\hat{\tau}_{tx}(t;x,\xi)) - \frac{Re(\hat{\tau}_{tt}(t;x,\xi))}{Im(\hat{\tau}_{tt}(t;x,\xi))}Im(\hat{\tau}_{tx}(t;x,\xi)). \qquad (3.75)
\end{aligned}
$$

We have the same conclusion for $\tilde{U}^y_q(s,\omega)$ and $\tilde{U}^x_\eta(s,\omega)$ as Lemma 3.3.1 and all terms involved are defined accordingly.

3.3.3 Asymptotic Analysis of Two Beams' Interaction

In this section, we would like to explore the interaction between two Gaussian beams, that is

$$
\iint drd\xi\bar{\hat{U}}^y_p(r,\omega)\hat{U}^x_\xi(r,\omega), \quad \iint dsd\eta\bar{\tilde{U}}^y_q(s,\omega)\tilde{U}^x_\eta(s,\omega).
$$

The rest of Section 3.3.3 is organized as the following. We will first discuss the distance between two beams $\hat{U}^y_p$ and $\hat{U}^x_\xi$ satisfying the parabolic scaling principle in the phase space. In Section 3.3.3.2, we will evaluate the error caused by replacing $\hat{U}^x_\xi(\cdot;\hat{t}_0(x,\xi))$ with $\hat{U}^x_\xi(\cdot;\hat{t}_0(y,p))$. After that, the difference between the exponents at different times is evaluated and it will allow us to map the phase term to the initial moment. In Section 3.3.3.4 and 3.3.3.5, we will compute the integral about r and show that there exists a Gaussian profile centered at $\hat{y}(\hat{t}_0(y,p))$. And the analysis can all be applied to the source side similarly.

Here we require that there is a significant interaction between two beams concerned, which means the distance between two central rays, $(\hat{y}(t),\hat{p}(t))$ and $(\hat{x}(t),\hat{\xi}(t))$ is less than the width of the beam $\hat{U}^y_p$ and the width of a beam is defined as $1/\sqrt{\epsilon}$, where ϵ is the smallest eigenvalue of $Im(\hat{M}(t))$.

3.3.3.1 Parabolic Scaling Principle

A wavepacket satisfying the so-called parabolic scaling principle means the wavelength of the typical oscillation of the wavepacket being equal to the square of the width of the wavepacket, and a Gaussian beam will satisfy parabolic scaling principle at any given time if it does initially. The following graph shows a single Gaussian wavepacket in $\mathbb{R}^2$ satisfying the parabolic scaling principle.

The following asymptotic analysis is needed throughout the proof.

Lemma 3.3.2. *Consider two scattering beams* $(\hat{y}(t),\hat{p}(t),\hat{M}(t),\hat{A}(t))$ *and*

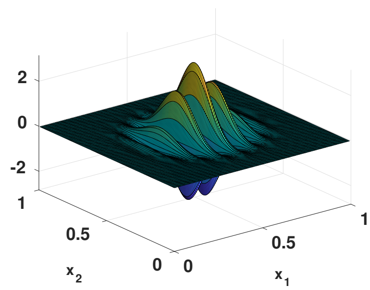


Figure 3.3: Real Part of Single Wavepacket $\phi_{l,i,k}$

$(\hat{x}(t), \hat{\xi}(t), \hat{N}(t), \hat{C}(t))$ *satisfying the parabolic scaling principle at the initial time, then*

$$||\hat{M}(t)|| \sim O(||p||), \quad ||\hat{N}(t)|| \sim O(||\xi||). \tag{3.76}$$

If there exists significant interaction effects between two beams, then

$$||\hat{y}(t) - \hat{x}(t)|| \sim O\left(\frac{1}{\sqrt{||p||}}\right). \tag{3.77}$$

Moreover, if we have $||p - \xi|| \sim O(\sqrt{||p||})$ *at the beginning,*

$$|\hat{\kappa}(t)| \sim 1 + O(||p||^{-\frac{1}{2}}), \quad ||\hat{\Xi}(t)|| \sim O(\sqrt{||p||}). \tag{3.78}$$

Proof. The assumption $||p - \xi|| \sim O(\sqrt{||p||})$ is a reasonable assumption, as we will see later $||p - \xi||$ will be controlled by a Gaussian profile, which means the value will be exponentially decaying when p and ξ are far from each other.

First, the Hamiltonian $G(x, p) = v(x)||p||$ remains as a constant along the ray. This implies that the order of the momentum $||p||$ will not change as the velocity v is a smooth function and bounded away from zero, that is $||\hat{p}(t)|| \sim ||p||$.

Second, we will evaluate $\hat{M}(t)$. The homogeneous of degree one Hamiltonian guarantees that Gaussian wavepackets satisfy the parabolic scaling principle at any given time. This implies that the size of Hessian $\hat{M}$ is around $O(||\hat{p}(t)||)$ and equation (3.76) is correct.

Due to the fact that $\hat{U}_p^y$ is well-localized around $\hat{y}(t)$ in the physical space, $\hat{U}_\xi^x$ will have small interaction with $\hat{U}_p^y$, if $\hat{x}(t)$ is beyond this localized region. And equation (3.77) is correct.

Third,

$$\begin{aligned}\left|\left|\frac{d(\hat{p}(t)-\hat{\xi}(t))}{dt}\right|\right| \leq & \left|\left|\nabla\nabla v(\hat{y}(t))(\hat{y}(t)-\hat{x}(t))||\hat{p}(t)||\right|\right| \\ & + ||\nabla v(\hat{y}(t))\left(\frac{\hat{p}(t)}{||\hat{p}(t)||}\right)^T(\hat{p}(t)-\hat{\xi}(t))|| \\ & \leq C_2||\hat{y}(t)-\hat{x}(t)||||\hat{p}(t)|| + C_1||\hat{p}(t)-\hat{\xi}(t)||,\end{aligned}$$

where C_1 and C_2 are the upper bound of $||\nabla v||$ and $||\nabla\nabla v||$ respectively. Moreover,

$$\frac{d||\hat{p}(t)-\hat{\xi}(t)||^2}{dt} = 2\frac{d(\hat{p}(t)-\hat{\xi}(t))}{dt}\cdot(\hat{p}(t)-\hat{\xi}(t)) \leq 2||\frac{d(\hat{p}(t)-\hat{\xi}(t))}{dt}||||\hat{p}(t)-\hat{\xi}(t)||,$$

by Cauchy-Schwartz inequality. We further get,

$$\begin{aligned}&\frac{d||\hat{p}(t)-\hat{\xi}(t)||^2}{dt} \leq C_1||\hat{p}(t)-\hat{\xi}(t)||^2 + C_2(||\hat{y}(t)-\hat{x}(t)||)(||\hat{p}(t)||)(||\hat{p}(t)-\hat{\xi}(t)||) \\ &\leq C_1||\hat{p}(t)-\hat{\xi}(t)||^2 + \frac{1}{2}(C_2||\hat{p}(t)||||\hat{y}(t)-\hat{x}(t)||)^2 + \frac{||\hat{p}(t)-\hat{\xi}(t)||^2}{2}.\end{aligned}$$

By Gronwall inequality, we have

$$||\hat{p}(t) - \hat{\xi}(t)||^2 \sim O(||p||),$$

since $C_2||\hat{p}(t)||||\hat{y}(t) - \hat{x}(t)||$ is uniformly bounded. Consequently, equation (3.78) is correct. $\square$

3.3.3.2 Difference between Two Interacted Beams' Traveltime

In this part, we would like to calibrate the beam $\hat{U}^x_\xi(r, \omega; \hat{t}_0(x, \xi))$ according to the beam $\hat{U}^y_p(r, \omega)$ by shifting time $\hat{t}_0(x, \xi)$ to time $\hat{t}_0(y, p)$.

By Corollary 3.3.1,

$$\begin{aligned}\hat{U}^x_\xi(r, \omega) = {} & e^{i\hat{\varrho}(r, \hat{t}_0(x,\xi); x,\xi)} e^{-i\omega \hat{t}_0(x,\xi)} \\ & e^{i\hat{\gamma}(\hat{t}_0(x,\xi);x,\xi)|\omega - \hat{\tau}_t(\hat{t}_0(x,\xi);x,\xi) - \hat{\vartheta}(\hat{t}_0(x,\xi);x,\xi)^T (r - \hat{x}(\hat{t}_0(x,\xi)))|^2} \\ & \sqrt{\frac{i2\pi}{\hat{\tau}_{tt}(\hat{t}_0(x,\xi); x, \xi)}} \hat{C}(\hat{t}_0(x, \xi)) e^{-||r - \hat{x}(\hat{t}_0(x,\xi))||^2 \frac{\hat{\mathcal{N}}(\hat{t}_0(x,\xi);x,\xi,N_0)}{2}}. \end{aligned} \tag{3.79}$$

The difference $|\hat{t}_0(x, \xi) - \hat{t}_0(y, p)|$ is around $O\left(\frac{1}{\sqrt{||p||}}\right)$. To see this, the following bound can be obtained using Lemma 3.3.2,

$$|\hat{y}_d(t) - \hat{x}_d(t)| \leq ||\hat{y}(t) - \hat{x}(t)|| \sim O(\frac{1}{\sqrt{||p||}}). \tag{3.80}$$

By Assumption 3.2.1, we have

$$|\hat{t}_0(x,\xi) - \hat{t}_0(y,p)| \leq \frac{|\hat{y}_d(t) - \hat{x}_d(t)|}{\mathfrak{b}} \sim O\left(\frac{1}{\sqrt{||p||}}\right). \tag{3.81}$$

The following proof is essentially comparing each term of equation (3.79) at two different times $\hat{t}_0(x,\xi)$ and $\hat{t}_0(y,p)$.

First, $e^{-i\omega\hat{t}_0(x,\xi)}$ becomes,

$$e^{-i\omega\hat{t}_0(x,\xi)} = e^{-i\omega\hat{t}_0(y,p)} e^{-i\omega(\hat{t}_0(x,\xi) - \hat{t}_0(y,p))}. \tag{3.82}$$

Second, we will discuss $e^{i\hat{\gamma}(\hat{t}_0(x,\xi);x,\xi)|\omega - \hat{\tau}_t(\hat{t}_0(x,\xi);x,\xi) - \hat{\vartheta}(\hat{t}_0(x,\xi);x,\xi)^T(r - \hat{x}(\hat{t}_0))|^2}$.

Proposition 3.3.3.

$$\begin{aligned}
&i\hat{\gamma}(\hat{t}_0(x,\xi);x,\xi)|\omega - \hat{\tau}_t(\hat{t}_0(x,\xi);x,\xi) - \hat{\vartheta}(\hat{t}_0(x,\xi);x,\xi)^T(r - \hat{x}(\hat{t}_0(x,\xi)))|^2 = \\
&i\hat{\gamma}(\hat{t}_0(y,p);x,\xi)|\omega - \hat{\tau}_t(\hat{t}_0(y,p);x,\xi) - \left(\hat{\vartheta}(\hat{t}_0(y,p);x,\xi)\right)^T(r - \hat{x}(\hat{t}_0(y,p)))|^2 \\
&+ O\left(\frac{1}{||p||}\right).
\end{aligned} \tag{3.83}$$

Proof. See Appendix A. □

Third, we will discuss $||r - \hat{x}(\hat{t}_0(x,\xi))||^2_{\hat{\mathcal{N}}}$.

Proposition 3.3.4.

$$||r - \hat{x}(\hat{t}_0(x,\xi))||^2_{\hat{\mathcal{N}}(\hat{t}_0(x,\xi))} = ||r - \hat{x}(\hat{t}_0(y,p))||^2_{\hat{\mathcal{N}}(\hat{t}_0(y,p))} + O\left(\frac{1}{\sqrt{||p||}}\right). \tag{3.84}$$

Proof. According to Lemma A.0.1, we have,

$$\hat{\mathcal{N}}(\hat{t}_0(x,\xi);x,\xi) = Im(\hat{N})(\hat{t}_0(x,\xi);x,\xi) - \frac{Im(\hat{N})\hat{\xi}\hat{\xi}^T Im(\hat{N})}{\hat{\xi}^T Im(\hat{N})\hat{\xi}}(\hat{t}_0(x,\xi);x,\xi)).$$

Then,

$$\begin{aligned}
&\hat{\mathcal{N}}(\hat{t}_0(x,\xi);x,\xi)\hat{\xi}(\hat{t}_0(x,\xi)) \\
&= Im(\hat{N})(\hat{t}_0(x,\xi);x,\xi)\hat{\xi}(\hat{t}_0(x,\xi)) - Im(\hat{N})(\hat{t}_0(x,\xi);x,\xi)\hat{\xi}(\hat{t}_0(x,\xi)) = 0.
\end{aligned} \tag{3.85}$$

This is also correct for any vector parallel to $\hat{\xi}(\hat{t}_0(x,\xi))$. On the other hand,

$$\begin{aligned}
r - \hat{x}(\hat{t}_0(x,\xi)) &= r - \hat{x}(\hat{t}_0(y,p)) + \hat{x}(\hat{t}_0(y,p)) - \hat{x}(\hat{t}_0(x,\xi)) \\
&= r - \hat{x}(\hat{t}_0(y,p)) + G_p^{\pm}(\hat{x}(\hat{t}_0(y,p)),\hat{\xi}(\hat{t}_0(y,p)))(\hat{t}_0(y,p) - \hat{t}_0(x,\xi)) \\
&= r - \hat{x}(\hat{t}_0(y,p)) \pm v(\hat{x}(\hat{t}_0(y,p)))(\hat{t}_0(y,p) - \hat{t}_0(x,\xi))\frac{\hat{\xi}(\hat{t}_0(x,\xi))}{||\hat{\xi}(\hat{t}_0(x,\xi))||}.
\end{aligned} \tag{3.86}$$

Consequently,

$$\begin{aligned}
||r - \hat{x}(\hat{t}_0(x,\xi))||^2_{\hat{\mathcal{N}}(\hat{t}_0(x,\xi))} &= ||r - \hat{x}(\hat{t}_0(y,p)) + \lambda\hat{\xi}(\hat{t}_0(x,\xi))||^2_{\hat{\mathcal{N}}(\hat{t}_0(x,\xi))} \\
&= ||r - \hat{x}(\hat{t}_0(y,p))||^2_{\hat{\mathcal{N}}(\hat{t}_0(y,p))} + ||r - \hat{x}(\hat{t}_0(y,p))||^2_{\hat{\mathcal{N}}(\hat{t}_0(x,\xi)) - \hat{\mathcal{N}}(\hat{t}_0(y,p))} \\
&= ||r - \hat{x}(\hat{t}_0(y,p))||^2_{\hat{\mathcal{N}}(\hat{t}_0(y,p))} + O\left(\frac{1}{\sqrt{||p||}}\right).
\end{aligned} \tag{3.87}$$

$\square$

Finally, about $\hat{\varrho}(r,\hat{t}_0(x,\xi);x,\xi)$ defined in equation (3.74). All other terms are at constant order, except for $\hat{\xi}(\hat{t}_0(x,\xi)) \cdot (r - \hat{x}(\hat{t}_0(x,\xi)))$,

$$\begin{aligned}
&\hat{\xi}(\hat{t}_0(x,\xi)) \cdot (r - \hat{x}(\hat{t}_0(x,\xi))) = \hat{\xi}(\hat{t}_0(y,p)) \cdot (r - \hat{x}(\hat{t}_0(x,\xi))) \\
&= \hat{\xi}(\hat{t}_0(y,p)) \cdot \big(r - \hat{x}(\hat{t}_0(y,p))\big) \pm v(\hat{x}(\hat{t}_0(y,p)))(\hat{t}_0(y,p) - \hat{t}_0(x,\xi))||\hat{\xi}(\hat{t}_0(y,p))||) \\
&= \hat{\xi}(\hat{t}_0(y,p)) \cdot \big(r - \hat{x}(\hat{t}_0(y,p))\big) \pm v||\hat{p}(\hat{t}_0(y,p))||(\hat{t}_0(y,p) - \hat{t}_0(x,\xi)) + O(1) \\
&= \hat{\xi}(\hat{t}_0(y,p)) \cdot (r - \hat{x}(\hat{t}_0(y,p)) - \hat{\tau}_t(\hat{t}_0(y,p);y,p)(\hat{t}_0(y,p) - \hat{t}_0(x,\xi)) + O(1).
\end{aligned} \tag{3.88}$$

The asymptotic analysis in the last step comes from Lemma 3.3.2, that is $||\hat{p}(t) - \hat{\xi}(t)|| \sim O(\sqrt{||p||})$.

To expedite the discussion, we will use the following notations:

$$\begin{aligned}
\hat{t}_c &= \hat{t}_0(y,p), \\
\Delta\hat{t}_0(x,\xi;y,p) &= \hat{t}_0(y,p) - \hat{t}_0(x,\xi), \\
\tilde{t}_c &= \tilde{t}_0(y,q), \\
\Delta\tilde{t}_0(x,\eta;y,q) &= \tilde{t}_0(y,q) - \tilde{t}_0(x,\eta).
\end{aligned}$$

Lemma 3.3.3. *By Proposition 3.3.3, Proposition 3.3.4 and equation (3.82) and (3.88),*

$$\begin{aligned}
&\hat{U}^x_\xi(r,\omega;\hat{t}_0(x,\xi)) = \sqrt{\frac{i2\pi}{\hat{\tau}_{tt}(\hat{t}_c;x,\xi)}}\hat{C}(\hat{t}_c)e^{i\hat{\varrho}(r,\hat{t}_c;x,\xi) - i\hat{\xi}(\hat{t}_0(x,\xi))\cdot(r-\hat{x}(\hat{t}_0(x,\xi)))}e^{-i\omega\hat{t}_c} \\
&e^{i(\omega-\hat{\tau}_t(\hat{t}_c;x,\xi))\Delta\hat{t}_0(x,\xi;y,p)}e^{i\hat{\xi}(\hat{t}_c)\cdot(r-\hat{x}(\hat{t}_c))}e^{i\hat{\gamma}(\hat{t}_c;x,\xi)|\omega-\hat{\tau}_t(\hat{t}_c;x,\xi)-\hat{\vartheta}(\hat{t}_c;x,\xi)^T(r-\hat{x}(\hat{t}_c;x,\xi))|^2} \\
&e^{-||r-\hat{x}(\hat{t}_c;x,\xi)||^2\frac{\hat{\mathcal{N}}(\hat{t}_c)}{2}} + O\left(\frac{1}{\sqrt{||p||}}\right).
\end{aligned} \tag{3.89}$$

Now, since two beams are using the same traveltime, we will abbreviate parameter $\hat{t}_0$

in $\hat{U}^x_\xi(r,\omega;\hat{t}_0(x,\xi))$. On the source side,

Corollary 3.3.2. *When we calibrate the beam $\tilde{U}^x_\eta(s,\omega;\tilde{t}_0(x,\eta))$ according to the beam $\tilde{U}^y_q(s,\omega)$ by shifting time $\tilde{t}_0(x,\eta)$ to time $\tilde{t}_c$, we have*

$$\tilde{U}^x_\eta(s,\omega;\tilde{t}_0(x,\eta)) = \sqrt{\frac{i2\pi}{\tilde{\tau}_{tt}(\tilde{t}_c;x,\eta)}}\tilde{C}(\tilde{t}_c)e^{i\tilde{\varrho}(s,\tilde{t}_c;x,\eta)-i\tilde{\eta}(\tilde{t}_0(x,\eta))\cdot(s-\tilde{x}(\tilde{t}_0(x,\eta)))}e^{-i\omega\tilde{t}_c}$$
$$e^{i(\omega-\tilde{\tau}_t(\tilde{t}_c;x,\eta))\Delta\tilde{t}_0(x,\eta;y,q)}e^{i\tilde{\eta}(\tilde{t}_c)\cdot(s-\tilde{x}(\tilde{t}_c))}$$
$$e^{i\tilde{\gamma}(\tilde{t}_c;x,\eta)|\omega-\tilde{\tau}_t(\tilde{t}_c;x,\eta)-\tilde{\vartheta}(\tilde{t}_c;x,\eta)^T(s-\tilde{x}(\tilde{t}_c;x,\eta))|^2}e^{-||s-\tilde{x}(\tilde{t}_c;x,\eta)||^2\frac{\tilde{N}(\tilde{t}_c)}{2}} + O\left(\frac{1}{\sqrt{||q||}}\right). \tag{3.90}$$

3.3.3.3 Difference between Two Interacted Beams' Phase and Hessians

We will compare the difference between $\hat{M}(t)-\hat{N}(t)$ in this section. We first have the following inequality

Proposition 3.3.5. *Consider two scattering beams $(\hat{y}(t),\hat{p}(t),\hat{M}(t),\hat{A}(t))$ and $(\hat{x}(t),\hat{\xi}(t),\hat{N}(t),\hat{C}(t))$, and assume that there exists significant interaction effects between these two beams. There exists two bounded positive constants C_1^* and C_2^* related to the background velocity, such that*

$$\frac{d||\hat{M}(t)-\hat{N}(t)||}{dt} \leq C_1^*\sqrt{||p||} + C_2^*||\hat{M}(t)-\hat{N}(t)||. \tag{3.91}$$

where $||\hat{M}(t)-\hat{N}(t)||$ is defined as the matrix norm induced by the vector 2-norm.

Proof. See Appendix A. □

Lemma 3.3.4. *Consider two scattering beams* $(\hat{y}(t), \hat{p}(t), \hat{M}(t), \hat{A}(t))$ *and* $(\hat{x}(t), \hat{\xi}(t), \hat{N}(t), \hat{C}(t))$*, and there exists significant interaction effects between these two beams. Then*

$$||\hat{M}(t) - \hat{N}(t)|| \sim O\left(\sqrt{||p||}\right), \quad \forall t \in [0, T], \tag{3.92}$$

Proof. First, $\hat{M}(t) - \hat{N}(t)$ is zero at $t = 0$, since they satisfy the same initial condition. Using Proposition 3.3.5 and the fact that the norm $||\hat{M}(t) - \hat{N}(t)||$ is positive and both C_1^* and C_2^* are positive,

$$||\hat{M}(t) - \hat{N}(t)|| \leq C_1^* \sqrt{||p(0)||} T + \int_0^t C_2^* ||\hat{M}(s) - \hat{N}(s)|| ds, \tag{3.93}$$

since the boundary data $D(r, s, t)$ is measured in the time interval $[0, T]$. With Gronwall inequality,

$$||\hat{M}(t) - \hat{N}(t)|| \leq C_1^* T e^{C_2^* t} \sqrt{||p||}. \tag{3.94}$$

□

Moreover, we have the same conclusion for other related terms,

Corollary 3.3.3.

$$||\hat{\mathcal{M}}(t) - \hat{\mathcal{N}}(t)|| \sim O(\sqrt{||p||}), \quad ||\tilde{\mathcal{M}}(t) - \tilde{\mathcal{N}}(t)|| \sim O(\sqrt{||q||});$$

$$||\hat{\tau}_{tx}(\hat{t}_c; y, p) - \hat{\tau}_{tx}(\hat{t}_c; x, \xi)|| \sim O(\sqrt{||p||}), \quad ||\tilde{\tau}_{tx}(\tilde{t}_c; y, q) - \hat{\tau}_{tx}(\tilde{t}_c; x, \eta)|| \sim O(\sqrt{||q||});$$

$$|\hat{\tau}_{tt}(\hat{t}_c; y, p) - \hat{\tau}_{tt}(\hat{t}_c; x, \xi)| \sim O(\sqrt{||p||}), \quad |\tilde{\tau}_{tt}(\tilde{t}_c; y, q) - \tilde{\tau}_{tt}(\tilde{t}_c; x, \eta)| \sim O(\sqrt{||q||}).$$

Consequently, by Lemma A.0.1, both $\hat{U}_p^y(r, \omega)$ and $\hat{U}_\xi^x(r, \omega)$ are well-localized along the boundary, and we have

$$\frac{1}{2}||r-\hat{x}(t)||^2_{\hat{\mathcal{M}}-\hat{\mathcal{N}}} \sim O(||p||^{-\frac{1}{2}}),$$

which will be used in equation (3.101) when we replace $\hat{\mathcal{N}}$ with $\hat{\mathcal{M}}$. Similarly,

$$\frac{1}{2}||s-\tilde{x}(t)||^2_{\tilde{\mathcal{M}}-\tilde{\mathcal{N}}} \sim O(||q||^{-\frac{1}{2}}).$$

Lemma 3.3.5. *Consider two scattering beams $(\hat{y}(t), \hat{p}(t), \hat{M}(t), \hat{A}(t))$ and $(\hat{x}(t), \hat{\xi}(t), \hat{N}(t), \hat{C}(t))$, and there exists significant interaction effects between these two beams. Suppose the function $g(t)$ is*

$$g(t) = \hat{p}(t)\cdot(\hat{y}(t)-\hat{x}(t)), \tag{3.95}$$

then we have

$$g(t) = g(0) + O(1), \tag{3.96}$$

and

$$g^{'}(t) = v(\hat{x}(t))\left(\frac{1}{2}\frac{||\hat{\Xi}(t)||^2}{\hat{\kappa}(t)^2||\hat{p}(t)||}\right). \tag{3.97}$$

Proof. See Appendix A. □

Lemma 3.3.6. *Consider two scattering beams $(\hat{y}(t), \hat{p}(t), \hat{M}(t), \hat{A}(t))$ and $(\hat{x}(t), \hat{\xi}(t), \hat{N}(t), \hat{C}(t))$, and there exists significant interaction effects between these two beams. Suppose the pure imaginary matrix $\hat{M}(0)$ has a symmetric positive definite* *imaginary part and is the initial condition of the Hessian for the beam, then*

$$(y-x)^T \hat{M}(0)(y-x) = (\hat{y}(t)-\hat{x}(t))^T \hat{M}(t)(\hat{y}(t)-\hat{x}(t)) + O(1). \tag{3.98}$$

Proof. See Appendix A. □

3.3.3.4 Approximation of Two Beams' Interaction

In this section, we will use the conclusion obtained in previous sections to get the explicit formula of $\iint d\xi dr \bar{\hat{U}}_p^y(r,\omega)\hat{U}_\xi^x(r,\omega)$. We first have the proposition below which will be used in approximation,

Proposition 3.3.6. *The real-valued phase terms, $\hat{\varrho}(r,\hat{t}_c;y,p)-\hat{\theta}(r,\hat{t}_c;y,p)$ and $\hat{\varrho}(r,\hat{t}_c;x,\xi)-\hat{\theta}(r,\hat{t}_c;x,\xi)$, can be ignored since they are constant order terms with respect to the large wavenumber $||\xi_{l,i}|| = ||p+q||$.*

Proof. See Appendix A. □

According to equation (3.89), (3.63),

$$
\begin{aligned}
&\iint d\xi dr \bar{\hat{U}}_p^y(r,\omega)\hat{U}_\xi^x(r,\omega) = e^{iO(1)} \\
&\iint dr d\xi e^{i\omega\hat{t}_c - i\omega\hat{t}_c} e^{-i\hat{\varrho}(r,\hat{t}_c;y,p)+i\hat{\theta}(r,\hat{t}_c;y,p)} e^{-i\hat{\theta}(r,\hat{t}_c;y,p)} e^{\overline{(i\hat{\beta})}|\omega-\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\zeta}(r-\hat{y}(\hat{t}_c))|^2} \\
&e^{i\hat{\varrho}(r,\hat{t}_c;x,\xi)-i\hat{\theta}(r,\hat{t}_c;x,\xi)} e^{i\hat{\theta}(r,\hat{t}_c;x,\xi)} e^{i\hat{\gamma}|\omega-\hat{\tau}_t(\hat{t}_c;x,\xi)-\hat{\vartheta}(r-\hat{x}(\hat{t}_c))|^2} e^{i(\omega-\hat{\tau}_t(\hat{t}_c;y,p))\Delta\hat{t}_0(x,\xi;y,p)} \\
&e^{-||r-\hat{x}(\hat{t}_c)||^2_{\frac{\hat{\mathcal{N}}(\hat{t}_c)}{2}}} e^{-||r-\hat{y}(\hat{t}_c)||^2_{\frac{\hat{\mathcal{M}}(\hat{t}_c)}{2}}} \\
&= e^{iO(1)} \int d\xi e^{-i\omega\hat{t}_c + i\omega\hat{t}_c} e^{-\frac{1}{4}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}} e^{i\frac{1}{2}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{Re(\hat{M})(\hat{t}_c)}} \\
&\times \int dr e^{\overline{\left(i\hat{\beta}\right)}|\omega-\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\zeta}^T(r-\hat{y}(\hat{t}_c))|^2} e^{i\hat{\gamma}|\omega-\hat{\tau}_t(\hat{t}_c;x,\xi)-\hat{\vartheta}^T(r-\hat{x}(\hat{t}_c))|^2} e^{i\hat{\chi}(r,\hat{t}_c;y,p,x,\xi)} \\
&\times e^{i\hat{\xi}(\hat{t}_c)\cdot(r-\hat{x}(\hat{t}_c))-i\hat{p}(\hat{t}_c)\cdot(r-\hat{y}(\hat{t}_c))} e^{i(\omega-\hat{\tau}_t(\hat{t}_c;y,p))\Delta\hat{t}_0(x,\xi;y,p)} \\
&\times e^{-\frac{1}{2}||r-\hat{y}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}} e^{-\frac{1}{2}||r-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{N}}(\hat{t}_c)}} e^{\frac{1}{4}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}}, \qquad (3.99)
\end{aligned}
$$

where

$$
\begin{aligned}
\hat{\chi}(r,\hat{t}_c;y,p,x,\xi) = &-\frac{1}{2}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{Re(\hat{M})(\hat{t}_c)} - \frac{1}{2}||r-\hat{y}(\hat{t}_c)||^2_{Re(\hat{M})(\hat{t}_c)} \\
&+ \frac{1}{2}||r-\hat{x}(\hat{t}_c)||^2_{Re(\hat{N})(\hat{t}_c)}. \qquad (3.100)
\end{aligned}
$$

Here we neglect some constant order real-valued phase terms by Proposition 3.3.6, which can be considered as a smooth residual.

Then we replace $\hat{\mathcal{N}}$ with $\hat{\mathcal{M}}$, and the inner integral becomes,

$$\int dr e^{\overline{(i\hat{\beta})}|\omega-\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\zeta}^T(r-\hat{y}(\hat{t}_c))|^2} e^{i\hat{\gamma}|\omega-\hat{\tau}_t(\hat{t}_c;x,\xi)-\hat{\vartheta}^T(r-\hat{x}(\hat{t}_c))|^2}$$
$$\times e^{i\hat{\xi}(\hat{t}_c)\cdot(r-\hat{x}(\hat{t}_c))-i\hat{p}(\hat{t}_c)\cdot(r-\hat{y}(\hat{t}_c))} e^{i\hat{\chi}} e^{-\frac{1}{2}||r-\hat{y}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}}$$
$$\times e^{-\frac{1}{2}||r-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{N}}(\hat{t}_c)}} e^{\frac{1}{4}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}} e^{i(\omega-\hat{\tau}_t(\hat{t}_c;y,p))\Delta\hat{t}_0(x,\xi;y,p)} \tag{3.101}$$
$$= e^{i(\omega-\hat{\tau}_t(\hat{t}_c;y,p))\Delta\hat{t}_0(x,\xi;y,p)} e^{i\hat{p}(\hat{t}_c)\cdot(\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c))} e^{-i\frac{\hat{p}(\hat{t}_c)-\hat{\xi}(\hat{t}_c)}{2}\cdot(\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c))}$$
$$\int e^{i\hat{\chi}} e^{i\frac{\hat{\xi}(\hat{t}_c)-\hat{p}(\hat{t}_c)}{2}\cdot(2r-\hat{x}(\hat{t}_c)-\hat{y}(\hat{t}_c))} e^{-\frac{1}{4}||2r-\hat{x}(\hat{t}_c)-\hat{y}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}}$$
$$e^{\overline{(i\hat{\beta})}|\omega-\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\zeta}^T(r-\hat{y}(\hat{t}_c))|^2} e^{i\hat{\gamma}|\omega-\hat{\tau}_t(\hat{t}_c;x,\xi)-\hat{\vartheta}^T(r-\hat{x}(\hat{t}_c))|^2} dr + \left(\frac{1}{\sqrt{||p||}}\right). \tag{3.102}$$

The difference of replacing Hessian matrix has been evaluated in Lemma 3.3.4 and its Corollary. Obviously, now the integral about the receiver variable r is well-defined. We denote its result as $\hat{B}(x,\xi,\omega;y,p)$,

$$\hat{B}(x,\xi,\omega;y,p) = \int e^{i\hat{\chi}} e^{i\frac{\hat{\xi}(\hat{t}_c)-\hat{p}(\hat{t}_c)}{2}\cdot(2r-\hat{x}(\hat{t}_c)-\hat{y}(\hat{t}_c))} e^{-\frac{1}{4}||2r-\hat{x}(\hat{t}_c)-\hat{y}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}}$$
$$\times e^{\overline{(i\hat{\beta})}|\omega-\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\zeta}^T(r-\hat{y}(\hat{t}_c))|^2} e^{i\hat{\gamma}|\omega-\hat{\tau}_t(\hat{t}_c;x,\xi)-\hat{\vartheta}^T(r-\hat{x}(\hat{t}_c))|^2} dr. \tag{3.103}$$

Similarly,

$$\tilde{B}(x,\eta,\omega;y,q) = \int e^{i\tilde{\chi}} e^{i\frac{\tilde{\eta}(\tilde{t}_c)-\tilde{q}(\tilde{t}_c)}{2}\cdot(2s-\tilde{x}(\tilde{t}_c)-\tilde{y}(\tilde{t}_c))} e^{-\frac{1}{4}||2s-\tilde{x}(\tilde{t}_c)-\tilde{y}(\tilde{t}_c)||^2_{\tilde{\mathcal{M}}(\tilde{t}_c)}}$$
$$\times e^{\overline{(i\tilde{\beta})}|\omega-\tilde{\tau}_t(\tilde{t}_c;y,q)-\tilde{\zeta}^T(s-\tilde{y}(\tilde{t}_c))|^2} e^{i\tilde{\gamma}|\omega-\tilde{\tau}_t(\tilde{t}_c;x,\eta)-\tilde{\vartheta}^T(s-\tilde{x}(\tilde{t}_c))|^2} ds. \tag{3.104}$$

We can see that the integral about r and s is accounted for in the computation of $\hat{B}$ and $\tilde{B}$.

3.3.3.5 Integral about Boundary Points r and s

In this section, we will evaluate $\hat{B}$ and $\tilde{B}$ defined in equation (3.103) and (3.104) to show that they are essentially Gaussian functions. Continuing from the expression (3.103) of $\hat{B}(x,\xi,\omega;y,p)$, we first simplify the exponent $\hat{\chi}(r,\hat{t}_c;y,p,x,\xi)$,

$$\begin{aligned}\hat{\chi} &= -\frac{1}{2}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{Re(\hat{M})(\hat{t}_c)} - \frac{1}{2}||r-\hat{y}(\hat{t}_c)||^2_{Re(\hat{M})(\hat{t}_c)} + \frac{1}{2}||r-\hat{x}(\hat{t}_c)||^2_{Re(\hat{M})(\hat{t}_c)} \\ &= -\frac{1}{2}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{Re(\hat{M})(\hat{t}_c)} + \frac{1}{2}(\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c))^T Re(\hat{M})(\hat{t}_c)(2r-\hat{x}(\hat{t}_c)-\hat{y}(\hat{t}_c)),\end{aligned} \tag{3.105}$$

so that we have

$$\begin{aligned}\hat{B}(x,\xi,\omega;y,p) &= e^{-i\frac{1}{2}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{Re(\hat{M})(\hat{t}_c)}} \\ &\int e^{-i\left(\frac{\hat{p}(\hat{t}_c)-\hat{\xi}(\hat{t}_c)}{2}-\frac{Re(\hat{M})(\hat{t}_c)}{2}(\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c))\right)\cdot(2r-\hat{x}(\hat{t}_c)-\hat{y}(\hat{t}_c))} \\ &\times e^{-\frac{1}{4}||2r-\hat{x}(\hat{t}_c)-\hat{y}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}}\hat{\mathcal{F}}(r,x,\xi,\omega;y,p)dr,\end{aligned} \tag{3.106}$$

where

$$\hat{\mathcal{F}}(r,x,\xi,\omega;y,p) = \\ e^{-Im(\hat{\beta})|\omega-\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\zeta}^T(r-\hat{y}(\hat{t}_c))|^2} e^{-Im(\hat{\gamma})|\omega-\hat{\tau}_t(\hat{t}_c;x,\xi)-\hat{\vartheta}^T(r-\hat{x}(\hat{t}_c))|^2} e^{i\hat{g}(r,x,\xi,\omega;y,p)}, \tag{3.107}$$

and

$$\hat{g}(r,x,\xi,\omega;y,p) = -Re(\hat{\beta})|\omega - \hat{\tau}_t(\hat{t}_c;y,p) - \hat{\zeta}^T(r-\hat{y}(\hat{t}_c))|^2$$
$$+ Re(\hat{\gamma})|\omega - \hat{\tau}_t(\hat{t}_c;x,\xi) - \hat{\vartheta}^T(r-\hat{x}(\hat{t}_c))|^2. \tag{3.108}$$

Proposition 3.3.7. *The sum of first two terms in the exponent of $\hat{\mathcal{F}}$ in (3.107) satisfy,*

$$\begin{aligned}
&- Im(\hat{\beta})|\omega - \hat{\tau}_t(\hat{t}_c;y,p) - \hat{\zeta}^T(r-\hat{y}(\hat{t}_c))|^2 - Im(\hat{\gamma})|\omega - \hat{\tau}_t(\hat{t}_c;x,\xi) - \hat{\vartheta}^T(r-\hat{x}(\hat{t}_c))|^2 \\
&= -||r - \frac{\hat{x}(\hat{t}_c)+\hat{y}(\hat{t}_c)}{2}||^2_{2Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T} - \frac{Im(\hat{\beta})}{2}||\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c)||^2_{\hat{\zeta}\hat{\zeta}^T} \\
&- Im(\hat{\beta})|\omega - \hat{\tau}_t(\hat{t}_c;y,p)|^2 - Im(\hat{\gamma})|\omega - \hat{\tau}_t(\hat{t}_c;x,\xi)|^2 + O\left(\frac{1}{\sqrt{||p||}}\right),
\end{aligned} \tag{3.109}$$

where $\hat{\beta}$, $\hat{\gamma}$, $\hat{\zeta}$ and $\hat{\vartheta}$ are all defined at $\hat{t}_c$. Similarly, $\hat{g}$ in equation (3.108) is an $O(1)$ term.

Proof. See Appendix A. $\square$

Lemma 3.3.7. *The integral $\hat{B}(x,\xi,\omega;y,p)$ defined in (3.106) has a Gaussian profile centered at $\hat{y}(\hat{t}_c)$ and $\hat{p}(\hat{t}_c)$,*

$$\begin{aligned}
\hat{B}(x,\xi,\omega;y,p) &= e^{-\frac{i}{2}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{Re(\hat{M})(\hat{t}_c) - iIm(\hat{\beta})\hat{\zeta}\hat{\zeta}^T}} \\
& e^{-Im(\hat{\beta})|\omega-\hat{\tau}_t(\hat{t}_c;y,p)|^2 - Im(\hat{\gamma})|\omega-\hat{\tau}_t(\hat{t}_c;x,\xi)|^2} \sqrt{\frac{(2\pi)^{d-1}}{\det(\hat{\mathcal{M}}(\hat{t}_c) + \frac{Im(\hat{\beta})}{2}\hat{\zeta}\hat{\zeta}^T)}} \\
& e^{-||\left(\hat{p}(\hat{t}_c) - \hat{\xi}(\hat{t}_c) - Re(\hat{M})(\hat{t}_c)(\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c))\right)||^2_{(\hat{\mathcal{M}}(\hat{t}_c)+\frac{1}{2}Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T)^{-1}}} e^{iO(1)},
\end{aligned}$$

where we ignore the real-valued phase term $\hat{g}$.

Proof. Using Proposition 3.3.7, the integral $\hat{B}$ in (3.106) becomes the Fourier transform of Gaussian functions, so that we have

$$\hat{B}(x,\xi,\omega;y,p) = e^{-i\frac{1}{2}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{Re(\hat{M})(\hat{t}_c)-iIm(\hat{\beta})\hat{\zeta}\hat{\zeta}^T}}$$
$$e^{-Im(\hat{\beta})|\omega-\hat{\tau}_t(\hat{t}_c;y,p)|^2-Im(\hat{\gamma})|\omega-\hat{\tau}_t(\hat{t}_c;x,\xi)|^2}\sqrt{\frac{(2\pi)^{d-1}}{\det(\hat{\mathcal{M}}(\hat{t}_c)+\frac{Im(\hat{\beta})}{2}\hat{\zeta}\hat{\zeta}^T)}}$$
$$e^{-||\left(\hat{p}(\hat{t}_c)-\hat{\xi}(\hat{t}_c)-Re(\hat{M})(\hat{t}_c)(\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c))\right)||^2_{(\hat{\mathcal{M}}(\hat{t}_c)+\frac{1}{2}Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T)^{-1}}}. \tag{3.110}$$

□

Corollary 3.3.4. *Similarly, on the source side,*

$$\tilde{B}(x,\eta,\omega;y,q) = e^{-i\frac{1}{2}||\tilde{y}(\tilde{t}_c)-\tilde{x}(\tilde{t}_c)||^2_{Re(\tilde{M})(\tilde{t}_c)-iIm(\tilde{\beta})\tilde{\zeta}\tilde{\zeta}^T}}$$
$$e^{-Im(\tilde{\beta})|\omega-\tilde{\tau}_t(\tilde{t}_c;y,q)|^2-Im(\tilde{\gamma})|\omega-\tilde{\tau}_t(\tilde{t}_c;x,\eta)|^2}\sqrt{\frac{(2\pi)^{d-1}}{\det(\tilde{\mathcal{M}})(\tilde{t}_c)+\frac{Im(\tilde{\beta})}{2}\tilde{\zeta}\tilde{\zeta}^T)}}$$
$$e^{-||\left(\tilde{q}(\tilde{t}_c)-\tilde{\eta}(\tilde{t}_c)-Re(\tilde{M})(\tilde{t}_c)(\tilde{y}(\tilde{t}_c)-\tilde{x}(\tilde{t}_c))\right)||^2_{(\tilde{\mathcal{M}}(\tilde{t}_c)+\frac{1}{2}Im(\tilde{\beta})\tilde{\zeta}\tilde{\zeta}^T)^{-1}}}e^{iO(1)}. \tag{3.111}$$

3.3.3.6 Conclusion of Two Beams' Interaction

To summarize,

$$\int dr\bar{\hat{U}}^y_p(r,\omega)\int d\xi\hat{U}^x_\xi(r,\omega) \approx \int d\xi e^{-\frac{1}{4}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}}e^{i\frac{1}{2}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{Re(\hat{M})(\hat{t}_c)}}$$
$$e^{i(\omega-\hat{\tau}_t(\hat{t}_c;y,p))\Delta\hat{t}_0(x,\xi;y,p)}e^{i\hat{p}(\hat{t}_c)\cdot(\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c))}e^{i\frac{\hat{\xi}(\hat{t}_c)-\hat{p}(\hat{t}_c)}{2}\cdot(\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c))}\hat{B}(x,\xi,\omega;y,p). \tag{3.112}$$

By Lemma 3.3.5 and Lemma 3.3.6, equation (3.112) reduces to

$$\iint dr d\xi \bar{\hat{U}}_p^y(r,\omega)\hat{U}_\xi^x(r,\omega) = e^{ip\cdot(y-x)}\hat{H}(x,\omega;y,p), \tag{3.113}$$

where

$$\hat{H}(x,\omega;y,p) = \int d\xi e^{i(\hat{\psi}_1(\hat{t}_c)-\hat{\psi}_1(0))}e^{i(\hat{\psi}_2(\hat{t}_c)-\hat{\psi}_2(0))}e^{-\frac{1}{4}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}}$$
$$e^{i\frac{1}{2}(\hat{\xi}(\hat{t}_c)-\hat{p}(\hat{t}_c))\cdot(\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c))}\hat{B}(x,\xi,\omega;y,p)e^{i(\omega-\hat{\tau}_t(\hat{t}_c;y,p))\Delta\hat{t}_0(x,\xi;y,p)}, \tag{3.114}$$

and

$$\hat{\psi}_1(t;x,\xi,y,p) = \hat{p}(t)\cdot(\hat{y}(t)-\hat{x}(t)); \quad \hat{\psi}_1(0;x,\xi,y,p) = p\cdot(y-x);$$
$$\hat{\psi}_2(t;x,\xi,y,p) = \frac{1}{2}||\hat{y}(t)-\hat{x}(t)||^2_{Re(\hat{M})(t)}; \quad \hat{\psi}_2(0;x,\xi,y,p) = \frac{1}{2}||y-x||^2_{Re(\hat{M})(0)}. \tag{3.115}$$

The extra term $e^{-i\frac{1}{4}||y-x||^2_{Re(\hat{M})(0)}} = 1$, since $\hat{M}(0)$ is a pure imaginary matrix in Gaussian wavepacket transform. We denote

$$\hat{L}(x,\xi,\hat{t}_c;y,p) = \sum_{i=1,2}\hat{\psi}_i(\hat{t}_c)-\hat{\psi}_i(0)-\frac{1}{2}(\hat{p}(\hat{t}_c)-\hat{\xi}(\hat{t}_c))\cdot(\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)). \tag{3.116}$$

Lemma 3.3.5 and Lemma 3.3.6 guarantees $\hat{L}$ is $O(1)$.

Lemma 3.3.8. *The receiver-side beam interaction reduces to,*

$$\iint dr d\xi \bar{\hat{U}}_p^y(r,\omega)\hat{U}_\xi^x(r,\omega) \approx e^{ip\cdot(y-x)}\hat{H}(x,\xi,\omega;y,p),$$

where

$$\hat{H}(x,\xi,\omega;y,p) = \int d\xi e^{i\hat{L}(x,\xi,\hat{t}_c;y,p)} e^{-\frac{1}{4}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}} e^{-i\frac{1}{2}||y-x||^2_{Re(\hat{M})(0)}}$$
$$\times \hat{B}(x,\xi,\omega;y,p) e^{i(\omega-\hat{\tau}_t(\hat{t}_c;y,p))\Delta\hat{t}_0(x,\xi;y,p)}.$$

Corollary 3.3.5. *The source-side beam interaction reduces to,*

$$\iint ds d\eta \bar{\tilde{U}}^y_q(s,\omega)\tilde{U}^x_\eta(s,\omega) \approx e^{iq\cdot(y-x)}\tilde{H}(x,\eta,\omega;y,q),$$

where

$$\tilde{H}(x,\eta,\omega;y,q) = \int d\eta e^{i\tilde{L}(x,\eta,\tilde{t}_c;y,q)} e^{-\frac{1}{4}||\tilde{y}(\tilde{t}_c)-\tilde{x}(\tilde{t}_c)||^2_{\tilde{\mathcal{M}}(\tilde{t}_c)}} e^{-i\frac{1}{2}||y-x||^2_{Re(\tilde{M})(0)}}$$
$$\times \tilde{B}(x,\eta,\omega;y,q) e^{i(\omega-\tilde{\tau}_t(\tilde{t}_c;y,q))\Delta\tilde{t}_0(x,\eta;y,q)},$$

where $\tilde{L}$ is defined accordingly.

3.3.4 Asymptotic Analysis of Four Beams' Interaction

Using Lemma 3.3.8 and Corollary 3.3.5, the left-hand side of equation (3.47) now becomes,

$$
\begin{aligned}
& -\iiiint d\omega dr ds d\xi d\eta \omega^2 \bar{\hat{U}}^y_p(r,\omega)\bar{\tilde{U}}^y_q(s,\omega)\hat{U}^x_\xi(r,\omega)\tilde{U}^x_\eta(s,\omega) \\
& \approx -e^{i(p+q)(y-x)}\hat{H}(x,\xi,\omega;y,p)\tilde{H}(x,\eta,\omega;y,q) \\
& = -e^{i(p+q)(y-x)} \\
& \iint d\xi d\eta e^{i\hat{L}(x,\xi,\hat{t}_c;y,p)+i\tilde{L}(x,\eta,\tilde{t}_c;y,q)} e^{-\frac{1}{4}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}} e^{-\frac{1}{4}||\tilde{y}(\tilde{t}_c)-\tilde{x}(\tilde{t}_c)||^2_{\tilde{\mathcal{M}}(\tilde{t}_c)}} \\
& \int \omega^2 \hat{B}\tilde{B} e^{i(\omega-\hat{\tau}_t(\hat{t}_c;y,p))\Delta\hat{t}_0(x,\xi;y,p)} e^{i(\omega-\tilde{\tau}_t(\hat{t}_c;y,q))\Delta\tilde{t}_0(x,\eta;y,q)} d\omega.
\end{aligned}
\tag{3.117}
$$

We will discuss the interaction between four beams in this subsection.

3.3.4.1 Integral about Wavenumber ω

We will evaluate the first layer of integral (3.117) about frequency ω in this subsection. Before that, we define a function $K(p,q,y)$,

$$
\begin{aligned}
K(p,q,y) \equiv & -\left(\frac{Im(\hat{\beta})\left(\hat{\tau}_t(\hat{t}_c;y,p)\right)+Im(\tilde{\beta})\left(\tilde{\tau}_t(\tilde{t}_c;y,q)\right)}{Im(\hat{\beta})+Im(\tilde{\beta})}\right)^2 \\
& e^{-2\frac{Im(\hat{\beta})Im(\tilde{\beta})(\hat{\tau}_t(\hat{t}_c;y,p)-\tilde{\tau}_t(\tilde{t}_c;y,q))^2}{Im(\hat{\beta})+Im(\tilde{\beta})}},
\end{aligned}
\tag{3.118}
$$

a function $\hat{\mathcal{B}}(x,\xi;y,p,q)$ on the receiver side,

$$\hat{\mathcal{B}} = e^{-\frac{|\Delta\hat{t}_0(x,\xi;y,p)|^2}{4Im(\hat{\beta}+\tilde{\beta})}} e^{-Im(\hat{\beta})|\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\tau}_t(\hat{t}_c;x,\xi)|^2} e^{-\frac{Im(\hat{\beta})}{2}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\zeta}\hat{\zeta}^T}}$$
$$\times \sqrt{\frac{(2\pi)^{d-1}}{\det(\hat{\mathcal{M}}(\hat{t}_c)+\frac{Im(\hat{\beta})}{2}\hat{\zeta}\hat{\zeta}^T)}}$$
$$e^{-\left|\left|\hat{p}(\hat{t}_c)-\hat{\xi}(\hat{t}_c)-Re(\hat{M})(\hat{t}_c)(\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c))\right|\right|^2_{\left(\hat{\mathcal{M}}(\hat{t}_c)+2Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T\right)^{-1}}}, \tag{3.119}$$

and a function $\tilde{\mathcal{B}}(x,\eta;y,p,q)$ on the source side,

$$\tilde{\mathcal{B}} = e^{-\frac{|\Delta\tilde{t}_0(x,\eta;y,q)|^2}{4Im(\hat{\beta}+\tilde{\beta})}} e^{-Im(\tilde{\beta})|\tilde{\tau}_t(\tilde{t}_c;y,q)-\tilde{\tau}_t(\tilde{t}_c;x,\eta)|^2} e^{-\frac{Im(\tilde{\beta})}{2}||\tilde{y}(\tilde{t}_c)-\tilde{x}(\tilde{t}_c)||^2_{\tilde{\zeta}\tilde{\zeta}^T}}$$
$$\times \sqrt{\frac{(2\pi)^{d-1}}{\det(\tilde{\mathcal{M}}(\tilde{t}_c)+\frac{Im(\tilde{\beta})}{2}\tilde{\zeta}\tilde{\zeta}^T)}}$$
$$e^{-\left|\left|\tilde{q}(\tilde{t}_c)-\tilde{\eta}(\tilde{t}_c)-Re(\tilde{M})(\tilde{t}_c)(\tilde{y}(\tilde{t}_c)-\tilde{x}(\tilde{t}_c))\right|\right|^2_{\left(\tilde{\mathcal{M}}(\tilde{t}_c)+2Im(\tilde{\beta})\tilde{\zeta}\tilde{\zeta}^T\right)^{-1}}}. \tag{3.120}$$

Lemma 3.3.9. *The result after taking the integral about ω can be approximated,*

$$-\int e^{i(\omega-\hat{\tau}_t(\hat{t}_c;y,p))\Delta\hat{t}_0(x,\xi;y,p)} e^{i(\omega-\tilde{\tau}_t(\hat{t}_c;y,q))\Delta\tilde{t}_0(x,\eta;y,q)} \omega^2 \hat{B}\tilde{B} d\omega$$
$$= e^{iO(1)} K(p,q,y) \hat{\mathcal{B}}(x,\xi;y,p,q) \tilde{\mathcal{B}}(x,\eta;y,p,q) e^{-\frac{i}{2}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{Re(\hat{M}(\hat{t}_c))}}$$
$$e^{-\frac{i}{2}||\tilde{y}(\tilde{t}_c)-\tilde{x}(\tilde{t}_c)||^2_{Re(\tilde{M}(\tilde{t}_c))}}, \tag{3.121}$$

where both $\hat{\mathcal{B}}$ and $\tilde{\mathcal{B}}$ have phase functions with pure imaginary part only.

Proof. See Appendix A. □

3.3.4.2 Integral about Momentum ξ and η: Evaluation of Real Part of Phase

We can define the following integral directly from equation (3.117),

$$\hat{\mathcal{H}}(x;y,p,q) = \int e^{i\hat{L}(x,\xi,\hat{t}_c;y,p)}$$
$$e^{-i\frac{1}{2}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{Re(\hat{M})(\hat{t}_c)}}\hat{\mathcal{B}}(x,\xi;y,p,q)e^{-\frac{1}{4}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}}d\xi,$$

where

$$\hat{L}(x,\xi,\hat{t}_c;y,p) = \sum_{i=1,2}\hat{\psi}_i(\hat{t}_c) - \hat{\psi}_i(0) - \frac{1}{2}(\hat{p}(\hat{t}_c) - \hat{\xi}(\hat{t}_c)) \cdot (\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c)),$$

and recall that

$$\begin{aligned} &\hat{\psi}_1(t;x,\xi,y,p) = \hat{p}(t)\cdot(\hat{y}(t)-\hat{x}(t)); \quad \hat{\psi}_1(0;x,\xi,y,p) = p\cdot(y-x); \\ &\hat{\psi}_2(t;x,\xi,y,p) = \frac{1}{2}||\hat{y}(t)-\hat{x}(t)||^2_{Re(\hat{M})(t)}; \quad \hat{\psi}_2(0;x,\xi,y,p) = \frac{1}{2}||y-x||^2_{Re(\hat{M})(0)}. \end{aligned} \tag{3.122}$$

Define functions $\hat{\phi}_1$ and $\hat{\phi}_2$ as derivatives of equation (3.122),

$$\hat{\phi}_1(t;x,\xi,y,p) = \frac{d\hat{\psi}_1(t;x,\xi,y,p)}{dt}, \hat{\phi}_2(t;x,\xi,y,p) = \frac{d\hat{\psi}_2(t;x,\xi,y,p)}{dt}. \tag{3.123}$$

In this subsection, we will explore the real part of the phase function in $\hat{\mathcal{H}}$, i.e., $\hat{L} - \frac{1}{2}||\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c)||^2_{Re(\hat{M})}$.

Proposition 3.3.8. *The function $\hat{\phi}_1$ satisfies*

$$\hat{\phi}_1(t;x,\xi,y,p) = \frac{1}{2}||\hat{p}(t) - \hat{\xi}(t)||^2_{G_{pp}} + O\left(\frac{1}{\sqrt{||p||}}\right). \tag{3.124}$$

Proof. See Appendix A. □

Define a $2d$ by $2d$ real-valued matrix $\hat{\mathcal{R}}$,

$$\hat{\mathcal{R}}(t;y,p) = \begin{bmatrix} 0, & -\frac{1}{4}I \\ -\frac{1}{4}I, & tG_{pp}(\hat{y}(t),\hat{p}(t)) \end{bmatrix}. \tag{3.125}$$

Proposition 3.3.9. $\hat{\mathcal{H}}(x,y,p,q)$ *satisfies*

$$\hat{\mathcal{H}}(x,y,p,q) = e^{iO(1)} \int d\xi \hat{\mathcal{B}}(x,\xi;y,p,q) e^{-\frac{1}{4}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{M}}}} e^{i||(\hat{y}-\hat{x},\hat{p}-\hat{\xi})(\frac{\hat{t}_c}{2})||^2_{\hat{\mathcal{R}}\left(\frac{\hat{t}_c}{2}\right)}}.$$

Proof. First,

$$\begin{aligned} \hat{L} - \frac{1}{2}||\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c)||^2_{Re(\hat{M})(\hat{t}_c)} = \\ \hat{\psi}_1(\hat{t}_c) - \hat{\psi}_1(0) - \hat{\psi}_2(0) - \frac{1}{2}(\hat{p}(\hat{t}_c) - \hat{\xi}(\hat{t}_c)) \cdot (\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c)) \\ = \hat{\psi}_1(\hat{t}_c) - \hat{\psi}_1(0) - \frac{1}{2}(\hat{p}(\hat{t}_c) - \hat{\xi}(\hat{t}_c)) \cdot (\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c)), \end{aligned} \tag{3.126}$$

since $Re(\hat{M})(0) = 0$.

We approximate $\hat{L}$ as,

$$\hat{\psi}_1(\hat{t}_c) - \hat{\psi}_1(0) = \int_0^{\hat{t}_c} \hat{\phi}_1(t)dt \approx \hat{\phi}_1\left(\frac{\hat{t}_c}{2}\right)\hat{t}_c. \tag{3.127}$$

Moreover, by using Lemma 3.3.5 twice,

$$(\hat{p}(\hat{t}_c)-\hat{\xi}(\hat{t}_c))\cdot(\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c))=\left(\hat{p}\left(\frac{\hat{t}_c}{2}\right)-\hat{\xi}\left(\frac{\hat{t}_c}{2}\right)\right)\cdot\left(\hat{y}\left(\frac{\hat{t}_c}{2}\right)-\hat{x}\left(\frac{\hat{t}_c}{2}\right)\right)+O(1). \tag{3.128}$$

To summarize,

$$\begin{aligned}
&\hat{L}-\frac{1}{2}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{Re(\hat{M})(\hat{t}_c)}\approx\\
&2\frac{\hat{t}_c}{2}\hat{\phi}_1\left(\frac{\hat{t}_c}{2}\right)-\frac{1}{2}\left(\hat{p}\left(\frac{\hat{t}_c}{2}\right)-\hat{\xi}\left(\frac{\hat{t}_c}{2}\right)\right)\cdot\left(\hat{y}\left(\frac{\hat{t}_c}{2}\right)-\hat{x}\left(\frac{\hat{t}_c}{2}\right)\right)\\
&=\frac{\hat{t}_c}{2}||\hat{p}\left(\frac{\hat{t}_c}{2}\right)-\hat{\xi}\left(\frac{\hat{t}_c}{2}\right)||^2_{G_{pp}}-\frac{1}{2}\left(\hat{p}\left(\frac{\hat{t}_c}{2}\right)-\hat{\xi}\left(\frac{\hat{t}_c}{2}\right)\right)\cdot\left(\hat{y}\left(\frac{\hat{t}_c}{2}\right)-\hat{x}\left(\frac{\hat{t}_c}{2}\right)\right).\\
&=||(\hat{y}-\hat{x},\hat{p}-\hat{\xi})||^2_{\hat{\mathcal{R}}(\frac{\hat{t}_c}{2})}
\end{aligned} \tag{3.129}$$

$\square$

Corollary 3.3.6.

$$\tilde{\mathcal{H}}(x,y,p,q)\approx\int d\eta\tilde{\mathcal{B}}(x,\eta;y,p,q)e^{-\frac{1}{4}||\tilde{y}(\tilde{t}_c)-\tilde{x}(\tilde{t}_c)||^2_{\tilde{\mathcal{M}}}}\exp\left(i||(\tilde{y}-\tilde{x},\tilde{q}-\tilde{\eta})(\frac{\tilde{t}_c}{2})||^2_{\tilde{\mathcal{R}}(\frac{\tilde{t}_c}{2})}\right).$$

All terms are defined accordingly.

Now the right hand side of equation (3.117) is equal to $e^{i(p+q)(y-x)}\hat{\mathcal{H}}\tilde{\mathcal{H}}$. Therefore, we will focus on $\hat{\mathcal{H}}$ and $\tilde{\mathcal{H}}$ next.

3.3.4.3 Integral about Momentum ξ and η: Evaluation of Imaginary Part of Phase

After seeing the real part of the phase function in Proposition 3.3.9 is a complete quadratic term, we will explore more about the imaginary part in this section. Similar to the real part, we will prove the imaginary part is a complete quadratic term as well as a non-degenerate quadratic term.

We start with rewriting $\hat{\mathcal{B}}(x,\xi;y,p,q)$ in (3.119),

$$\hat{\mathcal{B}}(x,\xi;y,p,q) = \sqrt{\frac{(2\pi)^{d-1}}{\det(\hat{\mathcal{M}}(\hat{t}_c) + \frac{Im(\hat{\beta})}{2}\hat{\zeta}\hat{\zeta}^T)}} e^{-||(\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c),\hat{p}(\hat{t}_c)-\hat{\xi}(\hat{t}_c))||^2_{\hat{\mathcal{I}}(\hat{t}_c;y,p)}}$$
$$e^{-\frac{|\Delta\hat{t}_0(x,\xi;y,p)|^2}{4Im(\hat{\beta}+\tilde{\beta})}} e^{-Im(\hat{\beta})|\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\tau}_t(\hat{t}_c;x,\xi)|^2}.$$

where $\hat{\mathcal{I}}(\hat{t}_c;y,p)$ is a symmetric matrix depending on the fixed beam's parameters (y,p).

$$\hat{\mathcal{I}}(\hat{t}_c;y,p) = \begin{bmatrix} -Re(\hat{M}) \\ I \end{bmatrix} (\hat{\mathcal{M}} + 2Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T)^{-1} \begin{bmatrix} -Re(\hat{M}) & I \end{bmatrix} + \begin{bmatrix} \frac{Im(\hat{\beta})}{2}\hat{\zeta}\hat{\zeta}^T, & 0 \\ 0, & 0 \end{bmatrix}. \tag{3.130}$$

Using Proposition 3.3.9, we have

Lemma 3.3.10.

$$\hat{\mathcal{H}}(x;y,p,q) = \sqrt{\frac{(2\pi)^{d-1}}{\det(\hat{\mathcal{M}}(\hat{t}_c) + \frac{Im(\hat{\beta})}{2}\hat{\zeta}\hat{\zeta}^T)}} \int d\xi e^{-\frac{1}{4}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}}$$
$$\times e^{-Im(\hat{\beta})|\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\tau}_t(\hat{t}_c;x,\xi)|^2} e^{-\frac{|\Delta\hat{t}_0(x,\xi;y,p)|^2}{4Im(\hat{\beta}+\tilde{\beta})}}$$
$$\times \exp\left(i\begin{bmatrix}\hat{y}(\frac{\hat{t}_c}{2})-\hat{x}(\frac{\hat{t}_c}{2}), & \hat{p}(\frac{\hat{t}_c}{2})-\hat{\xi}(\frac{\hat{t}_c}{2})\end{bmatrix}\hat{\mathcal{R}}\begin{bmatrix}\hat{y}(\frac{\hat{t}_c}{2})-\hat{x}(\frac{\hat{t}_c}{2})\\ \hat{p}(\frac{\hat{t}_c}{2})-\hat{\xi}(\frac{\hat{t}_c}{2})\end{bmatrix}\right)$$
$$\times \exp\left(-\begin{bmatrix}\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c), & \hat{p}(\hat{t}_c)-\hat{\xi}(\hat{t}_c)\end{bmatrix}\hat{\mathcal{I}}\begin{bmatrix}\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)\\ \hat{p}(\hat{t}_c)-\hat{\xi}(\hat{t}_c)\end{bmatrix}\right). \quad (3.131)$$

Before proving $\hat{\mathcal{H}}$ has a non-degenerate Gaussian profile, we need one extra assumption.

Transformation $\mathcal{U}$ between Two Phase Spaces: Suppose there is a transformation $\mathcal{U}$ between two phase spaces governed by the certain Hamiltonian flow. The bi-characteristic of the beam $\hat{U}^x_\xi$ is initially in the phase space $\mathcal{P}_1 = \{(x,\xi), x \in \mathbb{R}^d, \xi \in \mathbb{R}^d\}$, then after propagating to the surface, the bi-characteristic is in the phase space $\mathcal{P}_2 = \{((t,x_*),(\omega,\xi_*)), t,\omega \in \mathbb{R}, x_*,\xi_* \in \mathbb{R}^{d-1}\}$.

$$\mathcal{U}(x,\xi) = \left((\hat{t}_0(x,\xi), \hat{x}_*(\hat{t}_0(x,\xi))), \left(-G\left(\hat{x}(\hat{t}_0(x,\xi)), \hat{\xi}(\hat{t}_0(x,\xi))\right), \hat{\xi}_*(\hat{t}_0(x,\xi))\right)\right),$$
(3.132)

where $\hat{t}_0(x,\xi)$ is the hitting time defined before, and

$\hat{x}(\hat{t}_0(x,\xi)) = (\hat{x}_*(\hat{t}_0(x,\xi)), 0)$ is the corresponding hitting point on the boundary. $G(x,\xi)$ is the associated Hamiltonian for the central ray and $\hat{\xi}_*$ is the component of the ray direction corresponding to the tangential direction of the surface $\{x \in \mathbb{R}^d, x_d = 0\}$, that is

$$\hat{\xi}_*(\hat{t}_0(x,\xi)) = (\hat{\xi}_1(\hat{t}_0(x,\xi)), \cdots, \hat{\xi}_{d-1}(\hat{t}_0(x,\xi))). \tag{3.133}$$

The component $\hat{\xi}_d$ corresponding to the normal direction of the surface can be uniquely defined by $(-G, \hat{\xi}_*)$ according to eikonal equation (2.6), so the degree of freedom won't change. We need an extra assumption about the bi-characteristic,

Assumption 3.3.1. *$\mathcal{U}$ is invertible.*

The Gaussian profile is only related to the imaginary part of the phase function, therefore, we first ignore the term associated with $\hat{\mathcal{R}}$ in (3.131).

Lemma 3.3.11. *There exists a full-rank $2d$ by $2d$ S.P.D. matrix $\hat{\mathcal{E}}$, such that*

$$-\frac{1}{4}||\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)} - Im(\hat{\beta})|\hat{\tau}_t(\hat{t}_c; y, p) - \hat{\tau}_t(\hat{t}_c; x, \xi)|^2 - \frac{|\Delta\hat{t}_0(x,\xi;y,p)|^2}{4Im(\hat{\beta}+\tilde{\beta})}$$
$$-\begin{bmatrix}\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c), \hat{p}(\hat{t}_c) - \hat{\xi}(\hat{t}_c)\end{bmatrix}\hat{\mathcal{I}}\begin{bmatrix}\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c)\\ \hat{p}(\hat{t}_c) - \hat{\xi}(\hat{t}_c)\end{bmatrix} = -||(y-x, p-\xi)||^2_{\hat{\mathcal{E}}}. \tag{3.134}$$

Proof. There are three steps to justify the non-degenerate Gaussian profile. First, we will use some approximations to move the left-hand side of equation (3.134) to the phase space $\mathcal{P}_2$, since $\hat{x}(\hat{t}_c)$ is not on the boundary. Second, we will prove that equation (3.134) in the phase space $\mathcal{P}_2$ is non-degenerate. Finally, we will use the transformation defined in Assumption 3.3.1 to move the left-hand side term from the phase space $\mathcal{P}_2$ to $\mathcal{P}_1$.

Recall that $\hat{t}_c = \hat{t}_0(y,p)$,

$$
\begin{aligned}
&-\frac{1}{4}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)} - Im(\hat{\beta})|\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\tau}_t(\hat{t}_c;x,\xi)|^2 - \frac{|\Delta\hat{t}_0(x,\xi;y,p)|^2}{4Im(\hat{\beta}+\tilde{\beta})} \\
&-\left[\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c),\hat{p}(\hat{t}_c)-\hat{\xi}(\hat{t}_c)\right]\hat{\mathcal{I}}\begin{bmatrix}\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)\\ \hat{p}(\hat{t}_c)-\hat{\xi}(\hat{t}_c)\end{bmatrix} \\
&=-\frac{1}{4}||\hat{y}(\hat{t}_0(y,p))-\hat{x}(\hat{t}_0(x,\xi))||^2_{\hat{\mathcal{M}}(\hat{t}_0(y,p))} - Im(\hat{\beta})|\hat{\tau}_t(\hat{t}_0(y,p);y,p)-\hat{\tau}_t(\hat{t}_0(x,\xi);x,\xi)|^2 \\
&-\frac{|\Delta\hat{t}_0(x,\xi;y,p)|^2}{4Im(\hat{\beta}+\tilde{\beta})} - \left[\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c),\hat{p}(\hat{t}_c)-\hat{\xi}(\hat{t}_c)\right]\hat{\mathcal{I}}\begin{bmatrix}\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)\\ \hat{p}(\hat{t}_c)-\hat{\xi}(\hat{t}_c)\end{bmatrix} + O\left(\frac{1}{\sqrt{||p||}}\right)
\end{aligned}
\tag{3.135}
$$

$$
\begin{aligned}
&=-\frac{1}{4}||\hat{y}(\hat{t}_0(y,p))-\hat{x}(\hat{t}_0(x,\xi))||^2_{\hat{\mathcal{M}}(\hat{t}_0(y,p))} - \frac{|\Delta\hat{t}_0(x,\xi;y,p)|^2}{4Im(\hat{\beta}+\tilde{\beta})} \\
&- Im(\hat{\beta})|\hat{\tau}_t(\hat{t}_0(y,p);y,p)-\hat{\tau}_t(\hat{t}_0(x,\xi);x,\xi)|^2 \\
&-\left|\left|\left[\hat{y}(\hat{t}_0(y,p))-\hat{x}(\hat{t}_0(x,\xi))\pm\Delta\hat{t}_0 v\frac{\hat{\xi}}{||\hat{\xi}||},\hat{p}(\hat{t}_0(y,p))-\hat{\xi}(\hat{t}_0(x,\xi))\right]\right|\right|^2_{\hat{\mathcal{I}}} + O\left(\frac{1}{\sqrt{||p||}}\right),
\end{aligned}
\tag{3.136}
$$

where v in equation (3.136) is defined at $\hat{y}(\hat{t}_0(y,p))$. The first step (3.135) is due to the following derivation. Using Corollary 3.3.3,

$$||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)} = ||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{N}}(\hat{t}_c)} + O\left(\frac{1}{\sqrt{|p||}}\right).$$

Using Proposition 3.3.4,

$$||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\mathcal{N}}(\hat{t}_c)} = ||\hat{y}(\hat{t}_0(y,p))-\hat{x}(\hat{t}_0(x,\xi))||^2_{\hat{\mathcal{N}}(\hat{t}_c)} + O\left(\frac{1}{\sqrt{|p||}}\right),$$

since $\hat{y}(\hat{t}_0(y,p))$ is on the boundary. Use Corollary 3.3.3 again,

$$\begin{aligned}
&||\hat{y}(\hat{t}_0(y,p)) - \hat{x}(\hat{t}_0(x,\xi))||^2_{\hat{\mathcal{N}}(\hat{t}_0(y,p))} = \\
&||\hat{y}(\hat{t}_0(y,p)) - \hat{x}(\hat{t}_0(x,\xi))||^2_{\hat{\mathcal{M}}(\hat{t}_0(y,p))} + O\left(\frac{1}{\sqrt{|p||}}\right).
\end{aligned}$$

The second step (3.136) is due to Assumption 3.2.3 as the ray direction $\hat{\xi}$ remains as a constant near the boundary.

Now we will prove equation (3.136) is non-degenerate in $\mathcal{P}_2$.

If $\hat{U}_p^y$ and $\hat{U}_\xi^x$ have different hitting points $\hat{y}(\hat{t}_c) \neq \hat{x}(\hat{t}_0(x,\xi))$, then equation (3.136) is obviously negative. This is also true for the nonzero difference of travel time $\Delta\hat{t}_0$ and $\hat{\tau}_t(\hat{t}_0(y,p);y,p) - \hat{\tau}_t(\hat{t}_0(x,\xi);x,\xi)$.

If we have all difference mentioned above is zero and $\hat{p}-\hat{\xi}$ is nonzero along the tangential directions of the boundary, then equation (3.136) becomes,

$$||(0, \hat{p}(\hat{t}_0(y,p)) - \hat{\xi}(\hat{t}_0(x,\xi)))||^2_{\hat{\mathcal{I}}} = ||\hat{p}(\hat{t}_c) - \hat{\xi}(\hat{t}_c)||^2_{(\hat{\mathcal{M}}+2Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T)^{-1}} > 0, \tag{3.137}$$

since $\hat{\mathcal{M}}$ and its pseudo-inverse $\hat{\mathcal{M}}^{-1}$ are S.P.D. if restricted to the tangential directions by Lemma A.0.1.

Finally, due to the existence of $\hat{\mathcal{U}}^{-1}$, we can define a non-degenerate Gaussian profile about (x,ξ) accordingly, that is

$$\hat{\mathcal{E}} = \hat{\mathcal{U}}^{-T}\hat{\mathcal{C}}\hat{\mathcal{U}}^{-1}, \tag{3.138}$$

where $\hat{\mathcal{C}}$ makes the following term

$$|| \left[(\Delta \hat{t}_0, \hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_0(x,\xi))), \quad (\hat{\tau}_t(\hat{t}_c; y, p) - \hat{\tau}_t(\hat{t}_c; y, p), \hat{p}(\hat{t}_c) - \hat{\xi}(\hat{t}_0(x,\xi))) \right] ||^2_{\hat{\mathcal{C}}}$$

equal to the equation (3.136). $\square$

Remark 3.3.1. *Although the real part of the phase function in equation (3.131) is ignored in this part of the computation, it will not affect the above computation essentially, especially the existence of the Gaussian profile.*

3.3.4.4 Integral about Momentum ξ and η

After justifying the existence of a non-degenerate Gaussian profile centered at (y, p), the next question is how to calculate $\hat{\mathcal{E}}$ in Lemma 3.3.11 numerically. We first define a matrix $\hat{\mathcal{A}}(\hat{t}_c; y, p, q)$ depending only on fixed beam's parameters (y, p)

$$\hat{\mathcal{A}}(\hat{t}_c; y, p, q) = \hat{\mathcal{I}}(\hat{t}_c; y, p) + \begin{bmatrix} \frac{1}{4}\hat{\mathcal{M}}(\hat{t}_c) + \frac{||\hat{p}(\hat{t}_c)||^2 e_d e_d^T}{4\left(v(\hat{y}(\hat{t}_c))\hat{p}_d(\hat{t}_c)\right)^2 Im(\hat{\beta}+\tilde{\beta})}, 0 \\ 0, Im(\hat{\beta}) v^2(\hat{y}(\hat{t}_c)) \left(\frac{\hat{p}(\hat{t}_c)}{||\hat{p}(\hat{t}_c)||}\right)\left(\frac{\hat{p}(\hat{t}_c)}{||\hat{p}(\hat{t}_c)||}\right)^T \end{bmatrix}, \tag{3.139}$$

where $e_d = (0, \cdots, 0, 1) \in \mathbb{R}^d$.

Lemma 3.3.12. *There exists a S.P.D. 2d by 2d matrix $\hat{\mathcal{A}}$, such that*

$$||(y - x, p - \xi)||^2_{\hat{\mathcal{E}}} = ||(\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c), \hat{p}(\hat{t}_c) - \hat{\xi}(\hat{t}_c))||^2_{\hat{\mathcal{A}}}, \tag{3.140}$$

and $\hat{\mathcal{A}}$ is defined in equation (3.139).

Proof. Firstly, we would like to approximate $\Delta \hat{t}_0(x, \xi; y, p)$ and $\hat{\tau}_t(\hat{t}_c; y, p) - \hat{\tau}_t(\hat{t}_c; x, \xi)$.

By the definition of hitting time and Assumption 3.2.3,

$$\begin{aligned} 0 &= \hat{y}_d(\hat{t}_c) - \hat{x}_d(\hat{t}_0(x,\xi)) \\ &= \hat{y}_d(\hat{t}_c) - \hat{x}_d(\hat{t}_c) \pm v(\hat{y}(\hat{t}_c)) \frac{\hat{p}_d(\hat{t}_c)}{||\hat{p}(\hat{t}_c)||} \Delta\hat{t}(x,\xi;y,p). \end{aligned} \tag{3.141}$$

Here $\pm$ is determined by the sign of the Hamiltonian. To summarize,

$$-\frac{|\Delta\hat{t}_0(x,\xi;y,p)|^2}{4Im(\hat{\beta}+\tilde{\beta})} = -\frac{||\hat{p}(\hat{t}_c)||^2 |e_d^T(\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c))|^2}{4\left(v(\hat{y}(\hat{t}_c))\hat{p}_d(\hat{t}_c)\right)^2 Im(\hat{\beta}+\tilde{\beta})}. \tag{3.142}$$

Similarly,

$$-Im(\hat{\beta})|\hat{\tau}_t(\hat{t}_c;y,p) - \hat{\tau}_t(\hat{t}_c;x,\xi)|^2 = -Im(\hat{\beta}) v^2(\hat{y}(\hat{t}_c)) \Big| ||\hat{p}(\hat{t}_c)|| - ||\hat{\xi}(\hat{t}_c)|| \Big|^2. \tag{3.143}$$

Furthermore, it can be approximated by,

$$-Im(\hat{\beta})|\hat{\tau}_t(\hat{t}_c;y,p) - \hat{\tau}_t(\hat{t}_c;x,\xi)|^2 \approx -Im(\hat{\beta}) v^2(\hat{y}(\hat{t}_c)) \left| \frac{\hat{p}(\hat{t}_c)}{||\hat{p}(\hat{t}_c)||} \cdot (\hat{p} - \hat{\xi}) \right|^2. \tag{3.144}$$

The Lemma is proved. $\square$

The integrand (3.131) now is a quadratic term about $(\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c), \hat{p}(\hat{t}_c) - \hat{\xi}(\hat{t}_c))$. Furthermore, we have the following proposition.

Proposition 3.3.10. *There exists a linear map $\hat{\mathcal{J}}(\hat{t}_c;y,p)$, such that*

$$\begin{bmatrix} \hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c) \\ \hat{p}(\hat{t}_c) - \hat{\xi}(\hat{t}_c) \end{bmatrix} \approx \hat{\mathcal{J}}(\hat{t}_c;y,p) \begin{bmatrix} y - x \\ p - \xi \end{bmatrix}. \tag{3.145}$$

Proof. See Appendix A. $\square$

Similarly, the matrix $\tilde{\mathcal{A}}(\tilde{t}_c; y, p, q)$ and the map $\tilde{\mathcal{J}}(\tilde{t}_c; y, q)$ on the source side are defined accordingly.

Now $\hat{\mathcal{H}}$ becomes,

$$\hat{\mathcal{H}}(x; y, p, q) = \sqrt{\frac{(2\pi)^{d-1}}{\det(\hat{\mathcal{M}}(\hat{t}_c) + \frac{Im(\hat{\beta})}{2}\hat{\zeta}\hat{\zeta}^T)}}$$
$$\int d\xi e^{-||\hat{\mathcal{J}}(\hat{t}_c;y,p)(y-x,\xi-p)||^2_{\hat{\mathcal{A}}(\hat{t}_c;y,p,q)}} e^{i||\hat{\mathcal{J}}(\frac{\hat{t}_c}{2};y,p)(y-x,\xi-p)||^2_{\hat{\mathcal{R}}(\hat{t}_c;y,p)}}. \tag{3.146}$$

Therefore, $\hat{\mathcal{E}}$ under Assumption 3.2.3 is

$$\hat{\mathcal{E}} \approx \hat{\mathcal{J}}(\hat{t}_c; y, p)^T \hat{\mathcal{A}}(\hat{t}_c; y, p, q) \hat{\mathcal{J}}(\hat{t}_c; y, p). \tag{3.147}$$

The above equation provides an efficient way to approximate $\hat{\mathcal{E}}$. Now let's compute the integral about ξ. We suppose the S.P.D. matrix $\hat{\mathcal{E}}$,

$$\hat{\mathcal{E}} = \begin{bmatrix} \hat{\mathcal{E}}_{11}, & \hat{\mathcal{E}}_{12} \\ \hat{\mathcal{E}}_{12}^T, & \hat{\mathcal{E}}_{22} \end{bmatrix}. \tag{3.148}$$

Now we would like to show that there's a Gaussian profile about x centered at y,

$$
\begin{aligned}
&\int d\xi e^{-||(y-x,p-\xi)||^2_{\hat{\mathcal{E}}}} \\
&= e^{-||y-x||^2_{\hat{\mathcal{E}}_{11}}} e^{||y-x||^2_{\hat{\mathcal{E}}_{12}\hat{\mathcal{E}}_{22}^{-1}\hat{\mathcal{E}}_{12}^T}} \int d\xi e^{-||p-\xi+\hat{\mathcal{E}}_{22}^{-1}\hat{\mathcal{E}}_{12}^T(y-x)||^2_{\hat{\mathcal{E}}_{22}}} \\
&= e^{-||y-x||^2_{\hat{\mathcal{E}}_{11}}} e^{||y-x||^2_{\hat{\mathcal{E}}_{12}\hat{\mathcal{E}}_{22}^{-1}\hat{\mathcal{E}}_{12}^T}} \int d\xi e^{-||p-\xi||^2_{\hat{\mathcal{E}}_{22}}} \\
&= e^{-||y-x||^2_{\hat{\mathcal{E}}_{11}}} e^{||y-x||^2_{\hat{\mathcal{E}}_{12}\hat{\mathcal{E}}_{22}^{-1}\hat{\mathcal{E}}_{12}^T}} \int d\xi e^{-||(0,p-\xi)||^2_{\hat{\mathcal{E}}}}. \qquad (3.149)
\end{aligned}
$$

$\hat{\mathcal{E}}_{11} - \hat{\mathcal{E}}_{12}\hat{\mathcal{E}}_{22}^{-1}\hat{\mathcal{E}}_{12}^T$ is S.P.D by the fact that it is the Schur complement of $\hat{\mathcal{E}}_{22}$ in $\hat{\mathcal{E}}$. This displays that $\hat{\mathcal{H}}$ contains a Gaussian profile about $y - x$.

Together with equation (3.147), we have

$$
\begin{aligned}
&\int d\xi e^{-||(y-x,p-\xi)||^2_{\hat{\mathcal{E}}}} \\
&= e^{-||y-x||^2_{\hat{\mathcal{E}}_{11}}} e^{||y-x||^2_{\hat{\mathcal{E}}_{12}\hat{\mathcal{E}}_{22}^{-1}\hat{\mathcal{E}}_{12}^T}} \int d\xi e^{-||\hat{\mathcal{J}}(\hat{t}_c;y,p)(0,p-\xi)||^2_{\hat{\mathcal{A}}}}.
\end{aligned}
$$

Together with equation (3.146), we obtain a way to calculate $\hat{\mathcal{H}}(x; y, p, q)$. The real part of the phase function associated with $\hat{\mathcal{R}}$ will not affect the final result essentially and it can be compensated by a constant order phase term.

$$
\begin{aligned}
\hat{\mathcal{H}}(x; y, p, q) = e^{iO(1)} &\sqrt{\frac{(2\pi)^{d-1}}{\det(\hat{\mathcal{M}}(\hat{t}_c) + \frac{Im(\hat{\beta})}{2}\hat{\zeta}\hat{\zeta}^T)}} e^{\frac{i}{4}||y-x||^2_{\hat{M}(0)}} \\
&\int d\xi e^{i||\hat{\mathcal{J}}(\hat{t}_c;y,p)(0,p-\xi)||^2_{\hat{\mathcal{A}}}} e^{i||\hat{\mathcal{J}}(\frac{\hat{t}_c}{2};y,p)(y-x,p-\xi)||^2_{\hat{\mathcal{R}}(\hat{t}_c;y,p)}}. \qquad (3.150)
\end{aligned}
$$

Here we use $\frac{1}{4}M_0$ to approximate the Schur complement of $\hat{\mathcal{E}}$.

Lemma 3.3.13. *$\hat{\mathcal{H}}$ can be approximated by the following equation,*

$$\hat{\mathcal{H}}(x; y, p, q) \approx \hat{\mathcal{K}}(y, p, q) e^{\frac{i}{4}||y-x||^2_{\hat{M}(0)}}, \tag{3.151}$$

where

$$\hat{\mathcal{K}}(y, p, q) = \sqrt{\frac{(2\pi)^{d-1}}{\det(\hat{\mathcal{M}}(\hat{t}_c) + \frac{Im(\hat{\beta})}{2}\hat{\zeta}\hat{\zeta}^T)}}$$
$$\int d\xi e^{-||\hat{\mathcal{J}}(\hat{t}_c;y,p)(0,\xi-p)||^2_{\hat{\mathcal{A}}(\hat{t}_c;y,p,q)}} e^{i||\hat{\mathcal{J}}(\frac{\hat{t}_c}{2};y,p)(0,\xi-p)||^2_{\hat{\mathcal{R}}(\hat{t}_c;y,p)}}. \tag{3.152}$$

Similarly, $\tilde{\mathcal{H}}$ can be approximated by the following equation,

Corollary 3.3.7.

$$\tilde{\mathcal{H}}(x; y, p, q) \approx \tilde{\mathcal{K}}(y, p, q) e^{\frac{i}{4}||y-x||^2_{\tilde{M}(0)}} \tag{3.153}$$

where

$$\tilde{\mathcal{K}}(y, p, q) = \sqrt{\frac{(2\pi)^{d-1}}{\det(\tilde{\mathcal{M}}(\tilde{t}_c) + \frac{Im(\tilde{\beta})}{2}\tilde{\zeta}\tilde{\zeta}^T)}}$$
$$\int d\eta e^{-||\tilde{\mathcal{J}}(\tilde{t}_c;y,q)(0,\eta-q)||^2_{\tilde{\mathcal{A}}(\tilde{t}_c;y,p,q)}} e^{i||\tilde{\mathcal{J}}(\frac{\tilde{t}_c}{2};y,q)(0,\eta-q)||^2_{\tilde{\mathcal{R}}(\tilde{t}_c;y,q)}}. \tag{3.154}$$

The real part $\hat{\mathcal{R}}(\hat{t}_c; y, p)$ and $\tilde{\mathcal{R}}(\tilde{t}_c; y, q)$ can be compensated by a constant phase term and will not affect the result essentially.

Remark 3.3.2. *In equation (3.151), we essentially approximate the Schur complement $\hat{\mathcal{E}}_{11} - \hat{\mathcal{E}}_{12}\hat{\mathcal{E}}_{22}^{-1}\hat{\mathcal{E}}_{12}^T$ by $\frac{M_0}{4}$. However, it is costly to use the exact value (3.146) since* *we have to store all matrices $\hat{\mathcal{A}}(\hat{t}_c; y, p, q), \tilde{\mathcal{A}}(\tilde{t}_c; y, p, q)$ generated by different pairs of $\{(p, q) : p + q = \xi_{l,i}\}$ to apply the inverse Gaussian wavepacket transform (3.49). On the other hand, the difference between equation (3.151) and equation (3.146) will be at constant order guaranteed by Lemma 3.3.6.*

To summarize, the central direction and central point of the wavepacket will not be affected by approximation (3.151) and the width of wavepacket is at the same scale. The exact Hessian information (3.146) can be covered but it is costly to compute.

Theorem 3.3.1.

$$\begin{aligned}&\int -\omega^2 d\omega \int drds \bar{\hat{U}}_p^y(r,\omega)\bar{\tilde{U}}_q^y(s,\omega) \int d\xi d\eta \hat{U}_\xi^x(r,\omega)\tilde{U}_\eta^x(s,\omega) \approx \\ &K(p,q,y)\hat{\mathcal{K}}(y,p,q)\tilde{\mathcal{K}}(y,p,q)e^{i(p+q)\cdot(y-x)}e^{i\|y-x\|^2_{M_0}/2},\end{aligned} \tag{3.155}$$

where $K(p, q, y)$ is defined in equation (3.118), $\hat{\mathcal{K}}$ and $\tilde{\mathcal{K}}$ are defined in equation (3.152)-(3.154). The distance between $\hat{\tau}_t(\hat{t}_c; y, p)$ and $\tilde{\tau}_t(\tilde{t}_c; y, q)$ is controlled by $K(p, q, y)$.

3.4 Implementation of the Prestack Imaging Operator

By equations (3.46), (3.47) and (3.48), we conclude that the partial imaging function $I_{pq}(y, \omega)$ is related to the Gaussian wavepacket transform of $\frac{2\alpha}{v^2}$ centered at y in the direction $p+q$.

$$\int I_{pq}(y,\omega)d\omega = E(p,q,y)\int dx\frac{2\alpha}{v^2}e^{i(p+q)\cdot(y-x)}e^{-||y-x||^2_{M_0}/2}, \tag{3.156}$$

where $E(p,q,y) = \hat{\mathcal{K}}(y,p,q)\tilde{\mathcal{K}}(y,p,q)K(p,q,y)$. The numerical scheme to calculate $E(p,q,y)$ is given by equation (3.118) and equations (3.152)-(3.154). Therefore, this section will be devoted to illustrating how to compute the integral of the imaging function $I_{pq}(y,\omega)$ efficiently.

$$\int I_{pq}(y,\omega)d\omega = \int d\omega \int drds\bar{\tilde{U}}^y_q(s,\omega)\bar{\hat{U}}^y_p(r,\omega)D(r,s,\omega). \tag{3.157}$$

We start with the integral about the wavenumber ω. Using Corollary 3.3.1,

$$\hat{\Phi}(r,\hat{t}_c;y,p) = \hat{\tau}_t(\hat{t}_c;y,p) + \hat{\zeta}^T(r-\hat{y}(\hat{t}_c)),$$
$$\tilde{\Phi}(s,\tilde{t}_c;y,q) = \tilde{\tau}_t(\tilde{t}_c;y,q) + \tilde{\zeta}^T(s-\tilde{y}(\tilde{t}_c)).$$

By considering the terms containing ω in $\tilde{U}^y_q(s,\omega)$ and $\hat{U}^y_p(r,\omega)$ only,

$$\begin{aligned}&\int d\omega\bar{\tilde{U}}^y_q(s,\omega)\bar{\hat{U}}^y_p(r,\omega)D(r,s,\omega) = \\ &e^{iO(1)}\int D(r,s,\omega)e^{i\omega(\hat{t}_c+\tilde{t}_c)}e^{-Im(\hat{\beta})|\omega-\hat{\Phi}|^2}e^{-Im(\tilde{\beta})|\omega-\tilde{\Phi}|^2}d\omega.\end{aligned} \tag{3.158}$$

Here we neglect the real part of the exponent, $Re(\hat{\beta})|\omega-\hat{\Phi}|^2$ and $Re(\tilde{\beta})|\omega-\tilde{\Phi}|^2$, since they are constant order terms by Proposition 3.3.7. Consequently, they are small compared with the term $\omega(\hat{t}_c+\tilde{t}_c)$.

The integral now can be considered as an inverse Fourier transform about wavenumber ω,

$$\int D(r,s,\omega)e^{i\omega(\hat{t}_c+\tilde{t}_c)}e^{-Im(\hat{\beta})|\omega-\hat{\Phi}|^2}e^{-Im(\tilde{\beta})|\omega-\tilde{\Phi}|^2}d\omega$$
$$=$$
$$e^{-\frac{Im(\hat{\beta})Im(\tilde{\beta})(\tilde{\Phi}-\hat{\Phi})^2}{Im(\hat{\beta}+\tilde{\beta})}}e^{iS(\hat{t}_c+\tilde{t}_c)}\int D(r,s,\omega)e^{i(\omega-S)(\hat{t}_c+\tilde{t}_c)}e^{-Im(\hat{\beta}+\tilde{\beta})(\omega-S)^2}d\omega, \tag{3.159}$$

where

$$S(r,s,\hat{t}_c,\tilde{t}_c;y,p,q)=\frac{Im(\hat{\beta})\hat{\Phi}(r,\hat{t}_c;y,p)+Im(\tilde{\beta})\tilde{\Phi}(s,\tilde{t}_c;y,q)}{Im(\tilde{\beta}+\hat{\beta})}. \tag{3.160}$$

The integral in equation (3.159) is indeed a convolution,

$$e^{iS(\hat{t}_c+\tilde{t}_c)}e^{-\frac{Im(\hat{\beta})Im(\tilde{\beta})(\tilde{\Phi}-\hat{\Phi})^2}{Im(\hat{\beta}+\tilde{\beta})}}\int D(r,s,\omega)e^{i(\omega-S)(\hat{t}_c+\tilde{t}_c)}e^{-Im(\hat{\beta}+\tilde{\beta})(\omega-S)^2}d\omega$$
$$=\sqrt{\frac{\pi}{Im(\hat{\beta}+\tilde{\beta})}}e^{iS(\hat{t}_c+\tilde{t}_c)}$$
$$e^{-\frac{Im(\hat{\beta})Im(\tilde{\beta})(\tilde{\Phi}-\hat{\Phi})^2}{Im(\hat{\beta}+\tilde{\beta})}}\int e^{-\frac{h^2}{4Im(\hat{\beta}+\tilde{\beta})}}e^{iSh}D(r,s,\hat{t}_c+\tilde{t}_c-h)dh. \tag{3.161}$$

Notice the data D in the above formula is in the time domain, which means there is no need to apply the Fourier transform to the data at prior. Moreover, the convolution integral is conducted in a small range, i.e. $O\left(\frac{1}{\sqrt{|\hat{\tau}_{tt}(\hat{t}_c;y,p)|+|\hat{\tau}_{tt}(\tilde{t}_c;y,q)|}}\right)$.

Naturally, use Corollary 3.3.1 and equation (3.157)

$$\int I_{pq}(y,\omega)d\omega = \sqrt{\frac{\pi}{Im(\hat{\beta}+\tilde{\beta})}}\overline{\hat{A}(\hat{t}_c)\tilde{A}(\tilde{t}_c)}$$
$$\int drdse^{-||r-\hat{y}(\hat{t}_c)||^2_{\frac{\hat{\mathcal{M}}}{2}}}e^{-||s-\tilde{y}(\tilde{t}_c)||^2_{\frac{\tilde{\mathcal{M}}}{2}}}e^{-\frac{Im(\hat{\beta})Im(\tilde{\beta})(\hat{\Phi}-\tilde{\Phi})^2}{Im(\hat{\beta}+\tilde{\beta})}}e^{iS(\hat{t}_c+\tilde{t}_c)}e^{-i(\hat{\varrho}(r)+\tilde{\varrho}(s))}$$
$$\times 2\pi\sqrt{\frac{1}{\hat{\tau}_{tt}(\hat{t}_c;y,p)\tilde{\tau}_{tt}(\tilde{t}_c;y,q)}}\int e^{-\frac{h^2}{4Im(\hat{\beta}+\tilde{\beta})}}e^{iSh}D(r,s,\hat{t}_c+\tilde{t}_c-h)dh. \tag{3.162}$$

The integral about r and s is easy due to the existence of Gaussian profiles which will constrain the integral range. Therefore, the regular integration scheme is enough.

Here we notice that the value of the imaging function is controlled by $|\tilde{\Phi}-\hat{\Phi}| \sim |\tilde{\tau}_t(\tilde{t}_c;y,q)-\hat{\tau}_t(\hat{t}_c;y,p)|$. Therefore, an efficient way is needed to avoid computing pairs of beams with little illumination or imaging function $\int d\omega I_{pq}(y,\omega)$ (3.162) which is small. Therefore, we should select pairs of beams (y,p) and (y,q) such that

$$|\hat{\Phi}(r,\hat{t}_c)-\tilde{\Phi}(s,\tilde{t}_c)| \le 2\sqrt{||p||}+2\sqrt{||q||}. \tag{3.163}$$

Fortunately, the time derivative of the phase function does not change along the ray, which means we can estimate equation (3.163) without propagating beams.

3.5 Numerical Results

3.5.1 Approximation of Beams along the surface

The numerical examples in this subsection are provided to justify the approximation in Section 3.3.2. The following tests are done by comparing our approximation results to the results obtained by the numerical integration method.

The velocity used here is $v = 0.8 + 0.4x_2$ and the initial beam is initiated at the subsurface point $(0, 0.5)$. The initial momentum is set as $30\pi[\cos(0.1\pi), \sin(0.1\pi)]$ and the initial amplitude is $1 + i$.

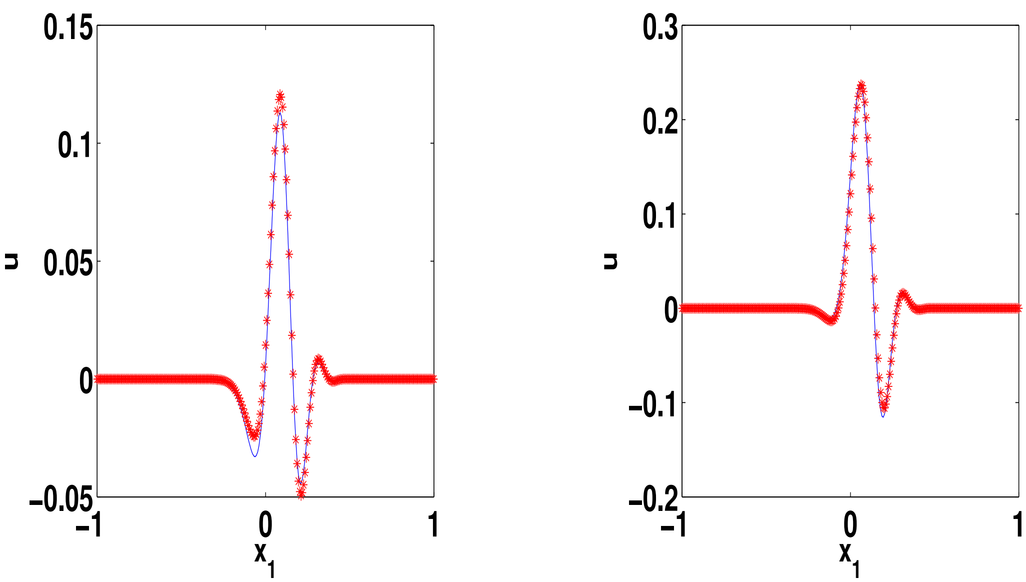


Figure 3.4: Numerical test for fast approximation of Gaussian beam along the surface. Left: ω = -110 Right: ω = -100.

The blue line represents the result achieved from the numerical integration method and the red star is the one from our fast approximation algorithm.

3.5.2 The Correctness of Prestack Imaging Operator

We conduct the numerical test to justify Theorem 3.3.1.

We first fix a direction $\boldsymbol{e}$ and the point y so that x is acquired by moving along this certain fixed direction, i.e. $x = y + \Delta h\boldsymbol{e}$. We compare our proposed result with the numerical integration result after removing the highly oscillated term $e^{ip\cdot(y-x)}$. The initial subsurface point is the point $y = (0, 0.5)$ and the initial ray direction is $p = (6\pi, -30\pi)$. Two different directions are picked here. The first one is $\boldsymbol{e} = (1, 1)$, while the second is chosen as $(1, 2)$. The x-axis in each plot is Δh , y-axis in each plot

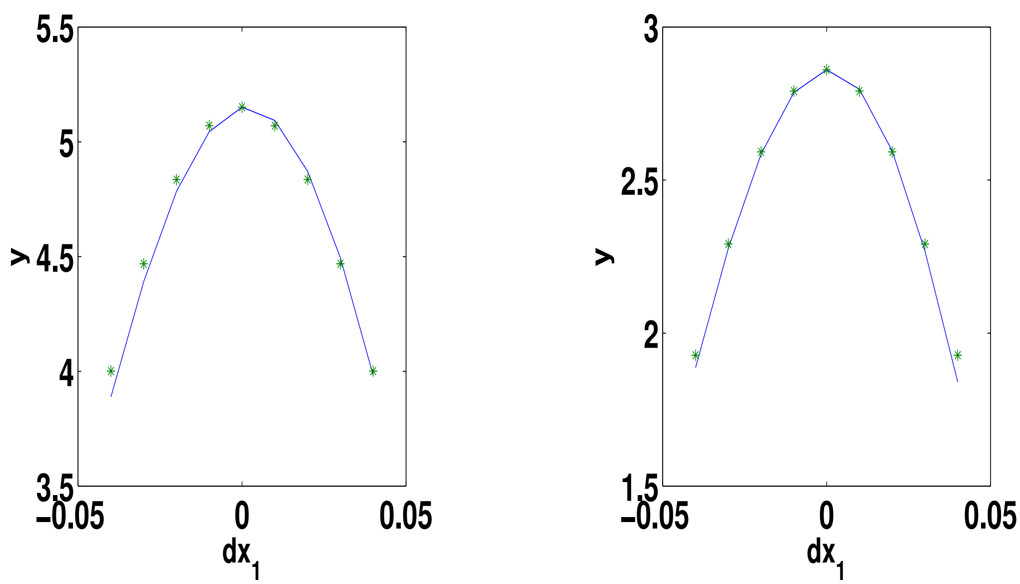
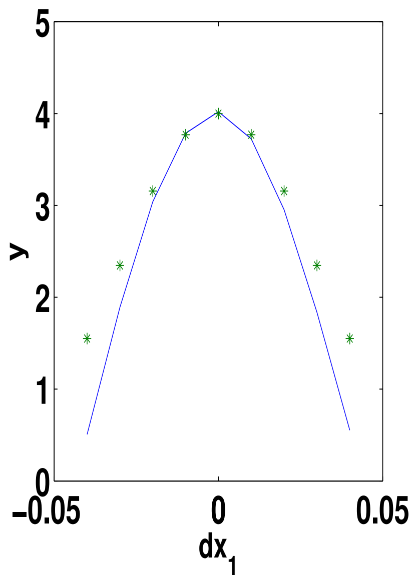


Figure 3.5: Constant Slowness: Theorem 3.3.1. Left: $\mathrm{d}x_1 = \mathrm{d}x_2$ Right: $\mathrm{d}x_1 = 2\mathrm{d}x_2$.

is the value of integral. As we can see in the constant slowness, the approximation proposed in Theorem 3.3.1 has only a small amount of error.

The second velocity used for test is $1 + 0.1x_2 + 0.1x_1$. Other setups are the same. As we can see in Fig. 3.6, the imaging operator does not perform as well as it does in the



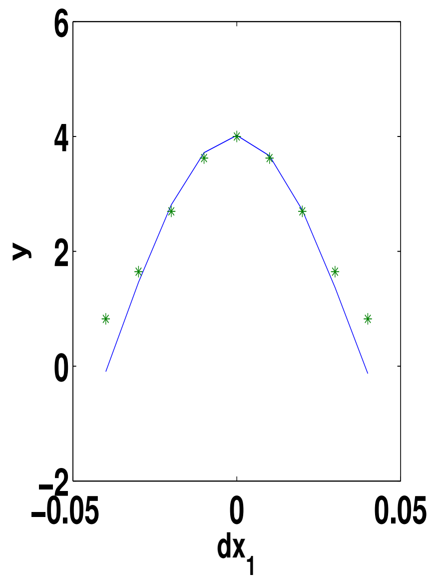


Figure 3.6: General Speed: Theorem 3.3.1. Left: $\mathrm{d}x_1 = \mathrm{d}x_2$ Right: $\mathrm{d}x_1 = 2\mathrm{d}x_2$.

constant slowness. However, the central momentum is captured correctly as stated in Remark 3.3.2.

3.5.3 Single Source Migration Test

We will recover the reflector by the single-source data trace in this section.

3.5.3.1 Example 1: Constant Background Slowness

Fig. 3.7 is the true slowness we employ, and there is a dipped layer. The source point here is at $x = (0, 0)$. As the Fig. 3.7 shows, the migration result shows the ability of our algorithm to detect the correct location and dipped angle.

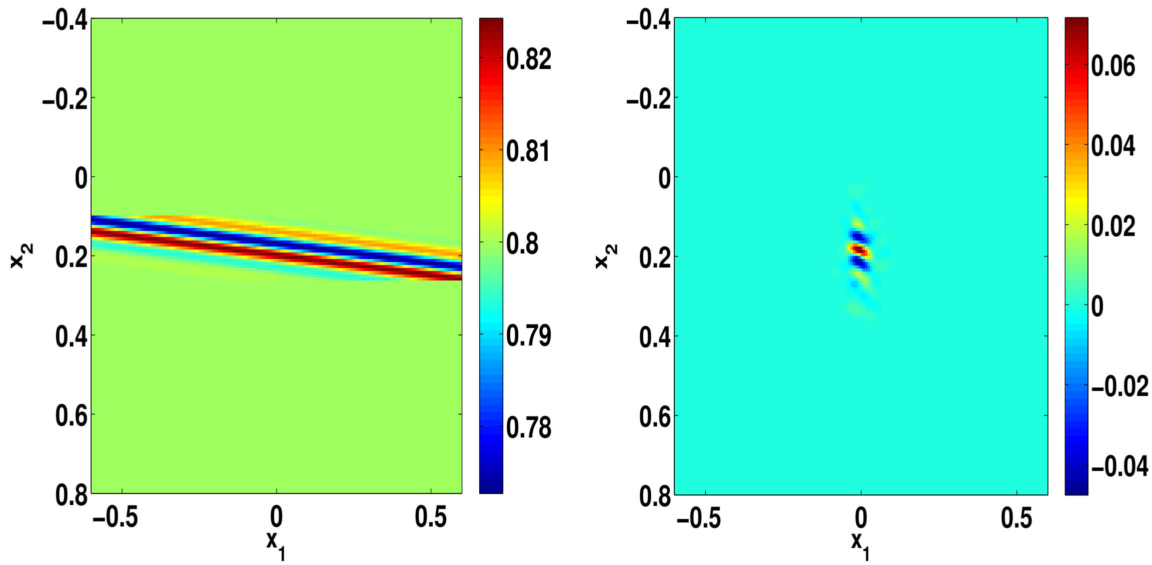


Figure 3.7: **Example 1:** Constant Slowness with Dipped layer. Left: True Slowness Right: Migration Result with single source trace.

3.5.3.2 Example 2: Multiple Flat Reflectors

In this numerical example, our migration algorithm is tested by two flat reflectors at different depth, The migration result is, The red dashed line is the true value while

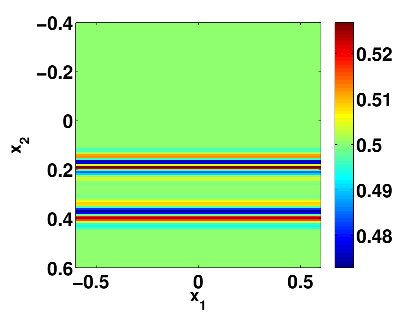


Figure 3.8: **Example 2:** True Slowness Model with Multiple Layers

the blue line the migration result. The deeper layer is not captured as well as the first layer. The error here is due to the Born approximation assumption.

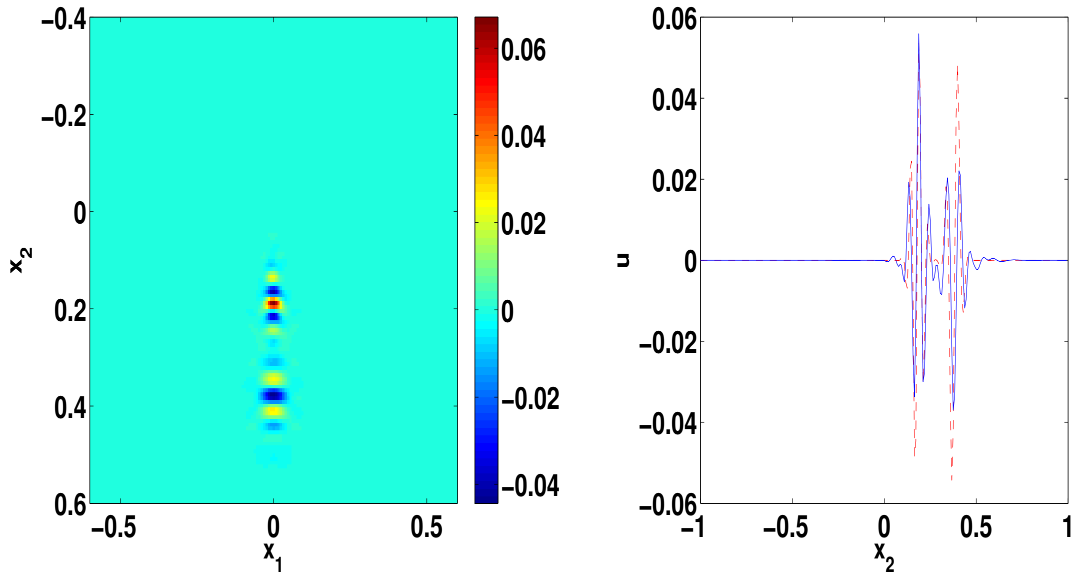


Figure 3.9: **Example 2:** Constant Slowness with Multiple Layers. Left: Migration Result over the Whole Space Right: Migration Result V.S. True Value at $x_1 = 0$.

3.5.3.3 Example 3: Linear Background Slowness

To see the amplitude information, we plot the slowness at $x_1 = 0$ and compare it with the true value. And the red dashed line is the true value while the blue line is our migration result.

3.5.4 Multiple Source Migration Test

We will use multiple-source data trace in this section.

3.5.4.1 Example 4: Constant Slowness with Dipped Layer

The background slowness is same as the one in Example 1. Sources are a series of points along the surface $-\frac{1}{4} : \frac{1}{16} : \frac{1}{4}$. This is applied to all multiple-source tests. The dipped layer is displayed correctly in Fig. 3.12.

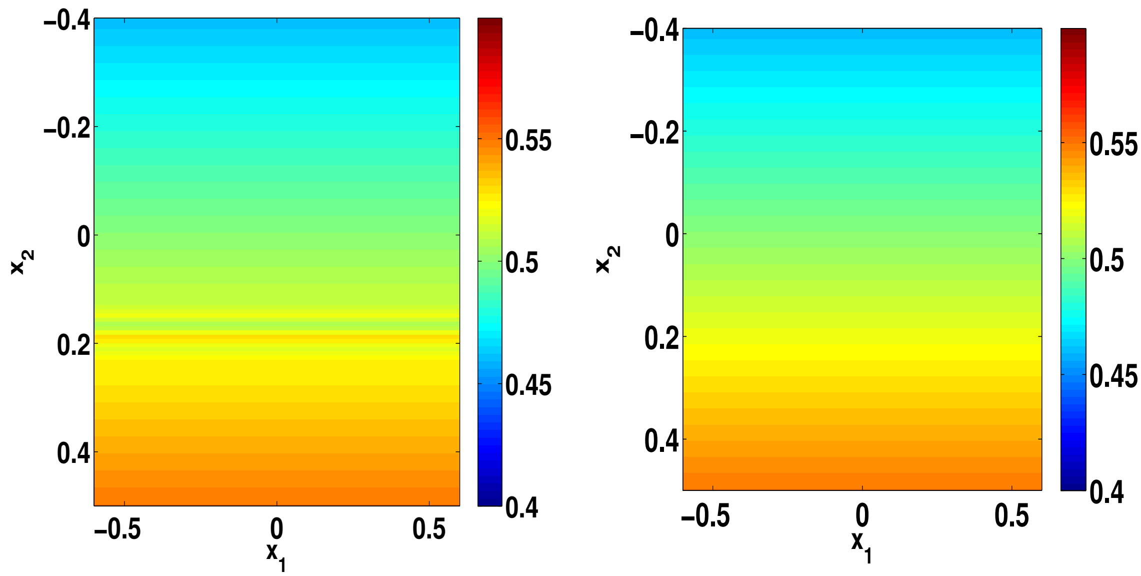


Figure 3.10: **Example 3:** Gradient Slowness Model. Left: True Slowness, Right: Smoothed Macro Slowness

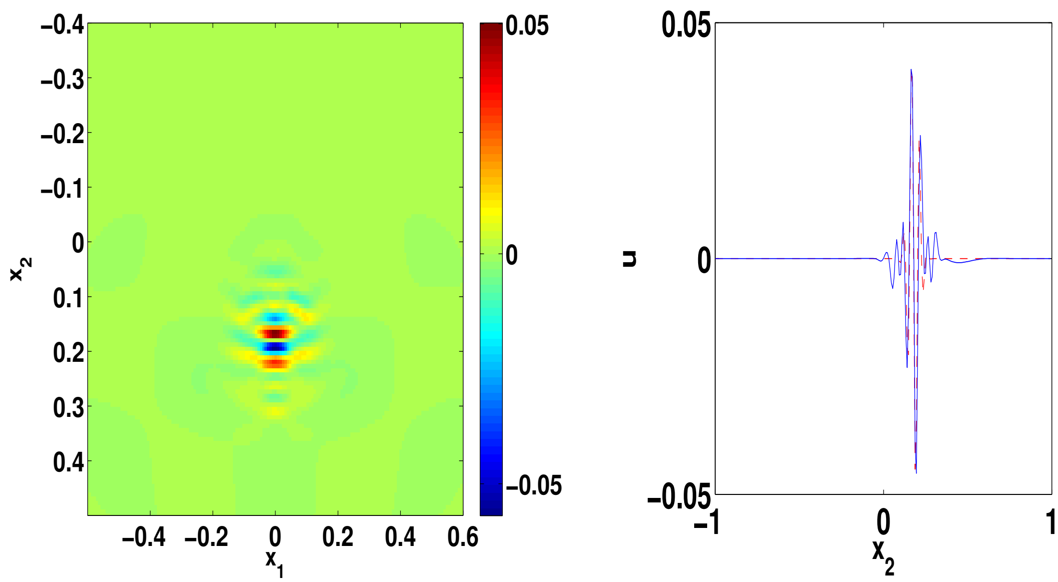


Figure 3.11: **Example 3:** Gradient Slowness Model. Left: Migration Result over the Whole Space, Right: Migration Result V.S. True Value at $x_1 = 0$.

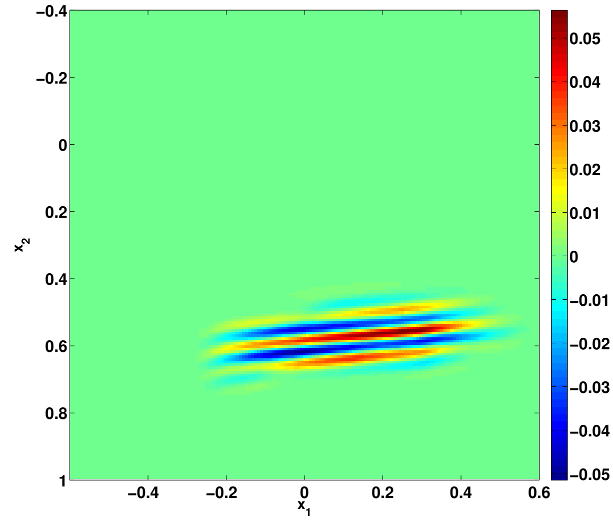


Figure 3.12: **Example 4:** Constant Slowness with the Dipped Layer (Multiple Sources)

3.5.4.2 Example 5: Flat Layer in Lateral Background Velocity

We add some lateral variation to the background slowness, i.e. $v = 0.8{+}0.1\sin(0.5\pi y)\sin(3\pi(x+0.05))$. The reflector is a horizontal reflector.

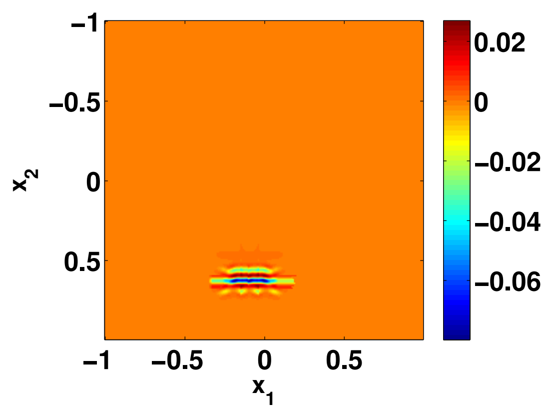


Figure 3.13: **Example 5:** Flat Layer in Lateral Background Velocity

3.5.4.3 Example 6: Slowness with Caustics

The next two examples in the section are both using the Gaussian velocity as the macro velocity shown in Fig. 3.14 (a). This is more complex as the caustics will show up. As we can see in Fig. 3.14, there is a caustics around the level $x_2 = 0.46$. Our

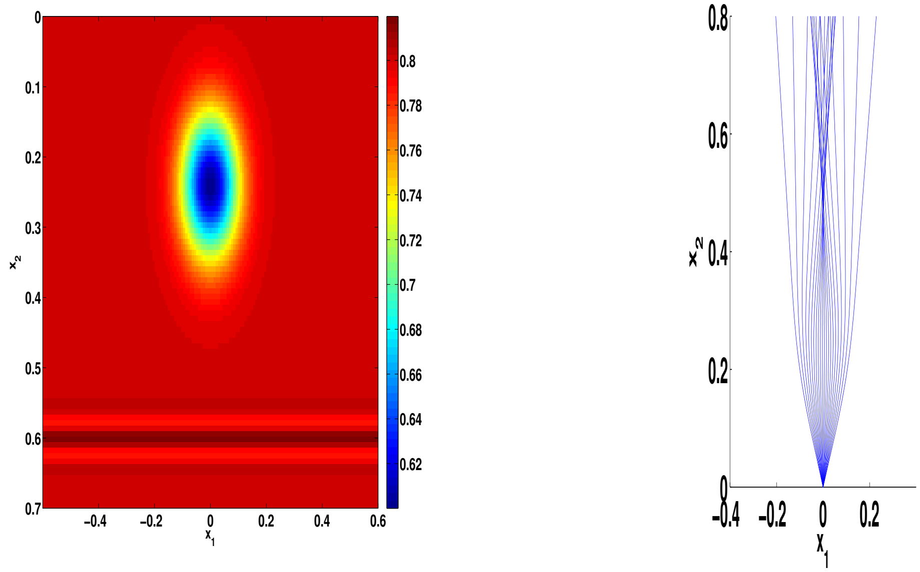


Figure 3.14: **Example 6:** Gaussian Slowness and its ray tracing. Left: Gaussian Slowness with Flat Reflector Right: Ray Tracing

flat reflector is below this caustics at $x_2 = 0.6$. The multi-value problem caused by caustics is resolved automatically by the Gaussian beam solution in Fig. 3.15.

3.5.4.4 Example 7: Polluted Trace Data

In the end, we would like to test our inversion process using the polluted data. We add 5% Gaussian noise into the synthetic data. See Fig. 3.16 for more details. The red dot line is the trace with extra Gaussian error, while the blue line is the original trace.

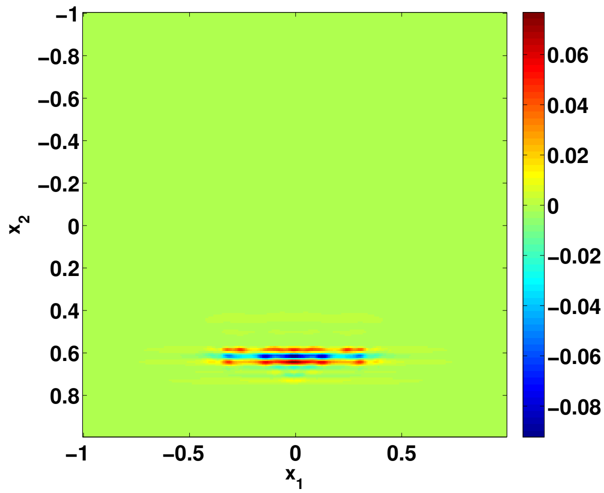


Figure 3.15: **Example 6:** Migration Result in the Gaussian Slowness with Caustics (Multiple Sources)

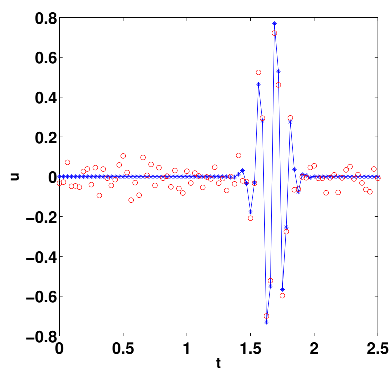


Figure 3.16: **Example 7:** True Trace V.S. Trace with Gaussian Error

The migration result is displayed in Fig. 3.17. There is no much difference in resulting images, especially around the reflector. To see more details, we compare two results at $x_2 = 0.65$ in Fig. 3.18, The red dot line comes from the non-polluted data while the blue line is from the polluted boundary data.

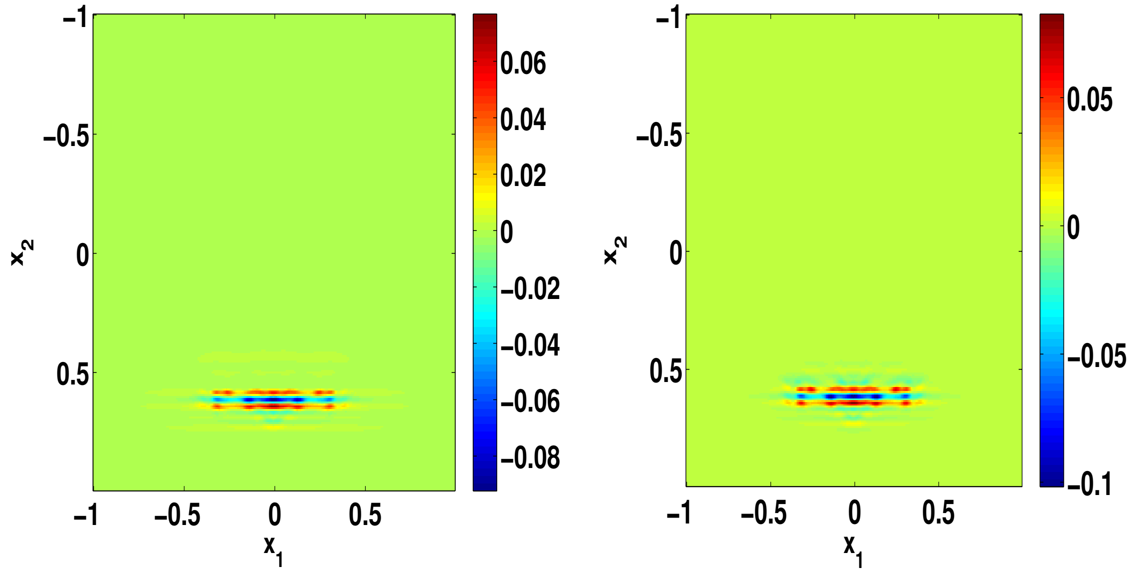


Figure 3.17: **Example 7:** Gaussian Slowness with polluted trace. Left: Migration result from Non-polluted Data; Right: Migration result from Polluted data

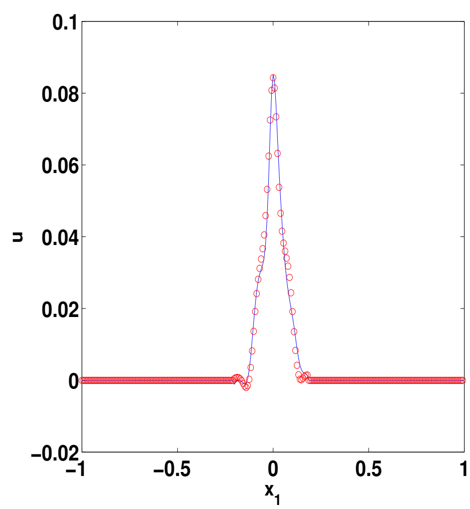


Figure 3.18: **Example 7:** Gaussian Slowness with polluted trace.Two Migration Results at $x_2 = 0.65$

Chapter 4

Fast Multiscale Gaussian Beam Method for Elastic Wave Equations in Bounded Domains

4.1 Asymptotic Method for the Elastic Wave equation

The problem considered in this paper is the initial-boundary value problem of the elastic wave equation.

$$0 = \rho\ddot{\boldsymbol{u}} - \nabla\lambda(\nabla\cdot\mathbf{u}) - \nabla\mu\cdot(\nabla\mathbf{u}+\nabla\mathbf{u}^T) - (\lambda+\mu)\nabla(\nabla\cdot\mathbf{u}) - \mu\Delta\mathbf{u}, \tag{4.1}$$

where the parameters λ and μ are known as the Lame parameters.

$$\dot{\boldsymbol{u}} = \frac{\partial\boldsymbol{u}}{\partial t}; \quad \ddot{\boldsymbol{u}} = \frac{\partial^2\boldsymbol{u}}{\partial t^2} \tag{4.2}$$

and this notation is applied to all functions. They are assumed to be smooth, positive and bounded away from zero. The initial condition is defined as the following,

$$\mathbf{u}(\mathbf{x},0) = \mathbf{f}; \quad \mathbf{u}_t(\mathbf{x},0) = \mathbf{g}, \tag{4.3}$$

where the functions $\mathbf{f}$ and $\mathbf{g}$ are compactly supported vector-valued functions in the space $L_2(R^d)$. We are looking for the asymptotic solution for the elastic wave equation (4.1) with two different types of boundary conditions, the periodic boundary condition and the homogeneous Dirichlet boundary condition, i.e.

$$\mathbf{u}(\mathbf{x},t)\Big|_{\partial\Omega\times[0,T]} = \mathbf{0}. \tag{4.4}$$

4.2 The Asymptotic Ansatz Solution to the Elastic Wave

We firstly derive the eikonal and the transport equation for the elastic wave equation. Same as the Geometrical-optics form, we consider the solution as the following series expansion,

$$\mathbf{u}(\mathbf{x},t) = e^{i\omega\tau(t,\mathbf{x})}\sum_{n=0}^{\infty}\mathbf{A}^{(n)}(t,\mathbf{x})(i\omega)^{-n}, \tag{4.5}$$

where the wavenumber ω is assumed to be a large parameter relative to the elastic moduli λ and μ's changing rate, i.e. $\frac{\omega L}{\min(\lambda,\mu)} >> 1$, where L is the characteristic distance defined as the scale over which the velocity changes slowly. The asymptotic solution for equation (4.1) is defined in the sense that both the equation itself and initial-boundary conditions are satisfied approximately with a small error when ω is large. To construct two equations governing the phase function τ and amplitude function $\mathbf{A}^{(0)}$ respectively, we substitute the ansatz form $\mathbf{A}^{(0)}e^{i\omega\tau}$ into equation (4.1). For convenience, we write $\mathbf{A}^{(0)}$ as $\mathbf{A}$. The j^{th} component of the elastic wave equation will then be

$$\begin{aligned} O\left(\frac{1}{\omega}\right) &= e^{i\omega\tau}\{\lambda_{,j}(A_{k,k} + i\omega A_k\tau_{,k}) + \mu_{,k}[A_{k,j} + A_{j,k} + i\omega(\tau_{,j}A_k + A_j\tau_{,k})] \\ &+ (\lambda+\mu)[\left(A_{k,k}\right)_{,j} + i\omega\left((A_k\tau_{,k})_{,j} + \tau_{,j}A_{k,k}\right) - \omega^2\tau_{,j}A_k\tau_{,k}] \\ &+ \mu[(A_{j,kk}) + i\omega(2A_{j,k}\tau_{,k} + A_j\tau_{,kk}) - A_j\omega^2\tau_{,k}^2] \\ &+ \rho A_j\omega^2(\dot{\tau})^2 - 2i\omega\rho\dot{\tau}\dot{A}_j - i\omega\rho A_j\ddot{\tau} - \rho\ddot{A}_j\}, \end{aligned} \tag{4.6}$$

here $|\cdot|$ denotes the Euclidean norm in R^d and

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_d} \end{pmatrix}.$$

Other notations used frequently in this paper are $\cdot$ representing the inner product between two column vectors and $\mathbf{v}^T$ representing the transpose of vector $\mathbf{v}$. Other notations we use in the above equation (4.6) are,

$$\tau_{,k} = \partial\tau/\partial x_k.$$

To make our derivation simpler, we let $\rho = 1$ without losing any generality.

4.2.1 P-wave and S-wave's Eikonal Equations

To cancel out the leading term ω^2 in equation (4.6), we set its coefficient to be zero,

$$0 = \dot{\tau}^2 A_j - \left((\lambda+\mu)\tau_{,j}A_k\tau_{,k} + \mu A_j\tau_{,k}\tau_{,k}\right). \tag{4.7}$$

After concatenating $j = 1, 2, \cdots, d$ as a vector, we will get

$$(\lambda+\mu)\nabla\tau\nabla\tau^T\mathbf{A} = (\dot{\tau}^2 - \mu\tau_{,k}\tau_{,k})\mathbf{A}, \tag{4.8}$$

so the amplitude vector $\mathbf{A}$ will be the eigenvector of the matrix $\nabla\tau\nabla\tau^T$. Notice that this simple matrix is rank one matrix by some linear algebra calculations, and the eigenvector corresponding to the single nonzero eigenvalue must be parallel to the vector $\nabla\tau^P$, while other two associated with the zero eigenvalues are orthogonal to $\nabla\tau^S$. The superscript here represents the category of their wave modes. We have

$$(\dot{\tau}^P)^2 - (\lambda+2\mu)(\tau^P_{,k}\tau^P_{,k}) = 0. \tag{4.9}$$

This is the eikonal equation for the P-wave whose amplitude vector $\mathbf{A}^P$ is parallel to the ray direction $\nabla\tau^P$. Another two eigenvectors are corresponding to the S-wave,

$$(\dot{\tau}^S)^2 - \mu(\tau^S_{,k}\tau^S_{,k}) = 0. \tag{4.10}$$

whose the amplitude vector $\mathbf{A}^S$ is perpendicular to the ray direction $\nabla\tau^S$.

With the eikonal equation at our disposal, we can apply the method of characteristics to the nonlinear eikonal equations (4.9) and (4.10). These two eikonal equations (4.9) and (4.10) are essentially the same, consequently, we consider them in the generic situation as a Hamilton-Jacobi equation.

$$\dot{\tau}^{P,S} + G^{P,S}(x, \nabla\tau) = 0, \tag{4.11}$$

where the Hamiltonian of the P-wave $G^P(x,p) = \pm\sqrt{\lambda + 2\mu}|p|$ and the Hamiltonian of the S-wave $G^S(x,p) = \pm\sqrt{\mu}|p|$. We consider the P-wave case for the illustration,

$$\begin{cases} \frac{dx}{dt} &= G_p^P(x(t), p(t)), \quad x(0) = x_0; \\ \frac{dp}{dt} &= -G_x^P(x(t), p(t)), \quad p(0) = p_0. \end{cases} \tag{4.12}$$

where t is the running parameter of a bicharacteristic.

Solving this ODE system yields a bicharacteristic in the phase space

$$\{(x(t), p(t)) : t \geq 0\}$$

and the associated ray $\gamma = \{x(t) : t \geq 0\}$, which is its x-component. Moreover, it is noticed that we have the equation $p(t) = \nabla\tau(t, x(t))$ along the ray γ due to the method of characteristics.

One of the most significant difference between the Gaussian beam and other ray-ansatz methods is that beams' phase functions are complex-valued. To be more specific, its second order derivative is complex-valued. To derive the dynamics of the Hessian matrix, we first differentiate the eikonal equation (4.11) with respect to t and x near the ray γ:

$$\dot{\tau}_x^P + G_x^P + \tau_{xx}^P G_p^P = 0, \tag{4.13}$$

$$\ddot{\tau}^P + (G_p^P)^T \dot{\tau}_x^P = 0. \tag{4.14}$$

Differentiating the first equation above (4.13) with respect to x again yields

$$\dot{\tau}_{xx}^P + G_{xx}^P + \tau_{xx}^P G_{xp}^P + (G_{xp}^P)^T \tau_{xx}^P + \tau_{xx}^P G_{pp}^P \tau_{xx}^p + \tau_{xxx}^P G_p^P = 0. \tag{4.15}$$

Since the equations (4.14), (4.13) and (4.15) are all valid everywhere in the phase space, it will still be valid if we concentrate them along the ray. Let $M^P(t)$ be the Hessian of the phase function along the ray

$$\frac{dM^P}{dt} + G_{xx}^P + M^P G_{xp}^P + (G_{xp}^P)^T M^P + M^P G_{pp}^P M^P = 0. \tag{4.16}$$

And the same rule can be applied to the S-wave with $G^S(x,p) = \pm\sqrt{\mu}|p|$. One interesting property of the Gaussian beam solution is that it will remain well-localized throughout the propagation, which means the imaginary part of the Hessian M should always be symmetric positive definite. The following lemma [49] guarantees this property throughout the propagation for all smooth ray trajectories,

Lemma 4.2.1. *If the Hamiltonian G is smooth enough, then the Hessian M(t) along the ray path γ has a positive-definite imaginary part, provided that it initially does.*

Accordingly, the Hessian of beam ansatz's phase functions is well-defined at all points even the caustics.

4.2.2 Transport Equation Governing P-wave's Amplitude Vectors

Setting the coefficient of $O(\omega)$ term in equation (4.6) equal to zero will yield the transport equation about the amplitude vector $\mathbf{A}$. We first see the j^{th} component,

$$\begin{aligned} 0 &= (2\dot{A}_j\dot{\tau} + \ddot{\tau}A_j) - \lambda_{,j}(A_k\tau_{,k}) - \mu_{,k}(\tau_{,j}A_k + A_j\tau_{,k}) \\ &- (\lambda+\mu)\left((A_k\tau_{,k})_{,j} + A_{k,k}\tau_{,j}\right) \\ &- \mu(2A_{j,k}\tau_{,k} + A_j\tau_{,kk}). \end{aligned} \tag{4.17}$$

Although it is complex at first glance, especially compared with the transport equation of the scalar wave equation, the complexity can be reduced by properties of the P-wave and S-wave.

We start with deriving the amplitude vector $\mathbf{A}^P$ for the P-wave, which is parallel to the ray direction $\nabla\tau^P$. Therefore, the P-wave's amplitude can be separated as $\mathbf{A} = a\nabla\tau^P$. To make derivation more readable, we write τ^P as τ in this part. We will yield the j^{th} component by inserting $\mathbf{A} = a\nabla\tau$ into equation (4.17)

$$\begin{aligned} 0 &= (2\dot{a}\tau_{,j}\dot{\tau} + 2a\dot{\tau}\dot{\tau}_{,j} + a\ddot{\tau}\tau_{,j}) - \lambda_{,j}a\tau_{,k}^2 - 2a(\mu_{,k}\tau_{,k})\tau_{,j} \\ &- (\lambda+\mu)\left(\tau_{,j}a_{,k}\tau_{,k} + a\tau_{,j}\tau_{,kk} + a_{,j}\tau_{,k}^2 + 2a\tau_{,k}\tau_{,kj}\right) \\ &- \mu\left(a\tau_{,j}\tau_{,kk} + 2\tau_{,j}a_{,k}\tau_{,k} + 2a\tau_{,k}\tau_{,kj}\right). \end{aligned}$$

Then we multiply the above equation with $\tau_{,j}$ and then sum over j,

$$(2\dot{a}\dot{\tau} + a\ddot{\tau})\tau_{,k}^2 + a\dot{\tau}|\nabla\tau|_t^2 = a|\nabla\tau|^2((\lambda+2\mu)_{,k}\tau_{,k}) + a(\lambda+2\mu)\tau_{,k}^2\tau_{,kk}$$
$$+ 2(\lambda+2\mu)\tau_{,k}^2(\tau_{,k}a_{,k}) + 2a(\lambda+2\mu)(\tau_{,j}\tau_{,jk}\tau_{,k}).$$

The term $a\dot{\tau}|\nabla\tau|_t^2$ on the left hand side of the above equation is equal to

$$\begin{aligned} a\dot{\tau}|\nabla\tau|_t^2 &= 2a\dot{\tau}(\tau_{,k}\dot{\tau}_{,k}) \\ &= a\tau_{,k}(\dot{\tau})_{,k}^2 \\ &= a\tau_{,k}((\lambda+2\mu)|\nabla\tau|^2)_{,k} \\ &= a|\nabla\tau|^2\tau_{,k}\,(\lambda+2\mu)_{,k} + 2a(\lambda+2\mu)\tau_{,j}\tau_{,jk}\tau_{,k}, \end{aligned} \tag{4.18}$$

since we are talking about the P-wave mode now and its phase τ satisfies the Hamiltonian-Jacobi equation $\dot{\tau}^2 - (\lambda+2\mu)|\nabla\tau|^2 = 0$. Consequently, the transport equation about the norm of the P-wave's amplitude vector $\mathbf{A}$ is

$$\dot{a} + \frac{(\lambda+2\mu)a_{,k}\tau_{,k}}{G} + \frac{a}{2G}((\lambda+2\mu)trace(M) - \ddot{\tau}) = 0. \tag{4.19}$$

Notice the second order derivatives of the phase function τ is involved, and the transport equation will be undefined if the phase function is not smooth. Lemma 4.2.1 guarantees a well-defined transport equation, while the classical Geometrical-Optical ansatz fails at the caustics region. Following [48], the ODE about the norm of amplitude can be added into the P-wave's dynamics by using equation (4.19) and $\frac{dx}{dt} = G_p$,

$$\frac{da(t, x(t))}{dt} + \frac{a}{2G}((\lambda + 2\mu)trace(M) - G_x \cdot G_p - G_p^T M G_p) = 0. \tag{4.20}$$

4.2.3 Transport Equation Governing S-wave's Amplitude Vectors

Now let's see the S-wave's case. We abbreviate τ^S as τ until the end of this section. Again, we first separate the amplitude vector as $\mathbf{A}^S = a\mathbf{D}$, where $\mathbf{D}$ is a vector which is orthogonal to the ray direction $\nabla\tau^S$ and its norm is fixed to be a constant. After substituting $\mathbf{A} = a\mathbf{D}$ into equation (4.17), we will have the following equation for the amplitude's j^{th} component,

$$\begin{aligned} 0 = & 2(\dot{a}D_j + a\dot{D}_j)\dot{\tau} + a\ddot{\tau}D_j - a(\mu_{,k}D_k)\tau_{,j} - a(\mu_{,k}\tau_{,k})D_j \\ & - (\lambda + \mu)((aD)_{k,k}\tau_{,j}) - 2\mu(aD_j)_{,k}\tau_{,k} - \mu a D_j \tau_{,kk}, \end{aligned} \tag{4.21}$$

for $j = 1, 2, 3$. After multiplying $\mathbf{D}_j$ with the above equation (4.21), we sum over the index j.

$$\left(a\mu_{,k}\tau_{,k} + a\mu\tau_{,kk} + 2\mu a_{,k}\tau_{,k}\right)|\mathbf{D}|^2 + 2\mu a\tau_{,k}|\boldsymbol{D}|^2_{,k} = a\ddot{\tau}|\mathbf{D}|^2 + 2\dot{a}\dot{\tau}|\mathbf{D}|^2 + a\dot{\tau}|\dot{\mathbf{D}}|^2, \tag{4.22}$$

since $\mathbf{D}$ is orthogonal to $\nabla\tau$. The last term in the above equation is zero as the norm of $\boldsymbol{D}$ is fixed, we have

$$(\nabla\tau \cdot \nabla)|\mathbf{D}|^2 = 0. \tag{4.23}$$

Therefore, the above equation can be simplified as,

$$a\mu_{,k}\tau_{,k} + a\mu\tau_{,kk} + 2\mu a_{,k}\tau_{,k} = 2\dot{a}\dot{\tau} + a\ddot{\tau}. \tag{4.24}$$

We fix the norm of $\mathbf{D}$ to be one for convenience and the above equation (4.24) provides the way to calculate the amplitude's norm a. It is not the same equation as the transport equation in the scalar wave equation. To yield the same equation, we divide $\sqrt{\mu}$ on both sides of equation (4.24),

$$\begin{aligned} \frac{2\dot{a}\dot{\tau} + a\ddot{\tau}}{\sqrt{\mu}} &= \frac{a}{\mu}\mu_{,k}\tau_{,k} + a\sqrt{\mu}\tau_{,kk} + 2\sqrt{\mu}a_{,k}\tau_{,k} \\ &= 2\tau_{,k}(\sqrt{\mu}a)_{,k} + (\sqrt{\mu}a)\tau_{,kk} \end{aligned}$$

If we set $\tilde{a} = \sqrt{\mu}a$ as new amplitude, then

$$2\dot{\tilde{a}}\dot{\tau} + \tilde{a}\ddot{\tau} = \mu(2\tau_{,k}\tilde{a}_{,k} + \tilde{a}\tau_{,kk}) \tag{4.25}$$

After yielding the same transport equation, we can obtain the same ODE as the P-wave (4.20).

Unlike the P-wave, we still need one more equation in the S-wave's ray system to describe the amplitude vector $\mathbf{A}^S$'s direction $\mathbf{D}$. To obtain the equation about the amplitude direction $\mathbf{D}$, we would like to plug equation (4.24) into equation (4.21) and the coeeficients in front of the direction $\boldsymbol{D}$ is zeros suggested by equation (4.24),

$$2a\dot{\tau}\dot{D}_j - 2\mu a(D_{j,i}\tau_{,i}) = \tau_{,j}\left(a\mu_{,i}D_i + (\lambda+\mu)(aD)_{k,k}\right) \tag{4.26}$$

The left hand side of equation (4.26) is equal to the term $\frac{dD_j(t,x(t))}{dt}$, since

$$\begin{aligned}
\dot{\tau}\frac{dD_j}{dt} &= \dot{\tau}\dot{D}_j + \dot{\tau}D_{j,i}\frac{dx_i}{dt} \\
&= \dot{\tau}\dot{D}_j - G(x(t),p(t))\boldsymbol{D}^T G_p(x(t),p(t)) \\
&= \dot{\tau}\dot{D}_j - \mu D_{j,i}\tau_{,i}(t,x(t))
\end{aligned} \tag{4.27}$$

We know that the left hand side of equation (4.26) is $2a\dot{\tau}\frac{dD_j}{dt}$, and the right hand side of equation (4.26) is parallel to $\nabla\tau$. Therefore, $\frac{d\boldsymbol{D}}{dt}$ is parallel to $\nabla\tau$. Together with the fact that the amplitude's direction $\mathbf{D}$ is always perpendicular to the ray direction $p(t) = \nabla\tau(t,x(t))$,

$$\begin{aligned}
0 &= \frac{dD_k p_k(t)}{dt} \\
0 &= \frac{dD_k}{dt}p_k(t) + \frac{dp_k(t)}{dt}D_k \\
\frac{dD_k}{dt} &= -\left(\frac{dp_j(t)}{dt}D_j\right)\frac{p_k(t)}{|p(t)|^2}
\end{aligned} \tag{4.28}$$

4.2.4 Single Beam Solution for P and S-wave

To summarize the ODE dynamics generated by the method of characteristics, we have

$$\begin{aligned}
\frac{d\boldsymbol{x}}{dt} &= G_p(\boldsymbol{x}(t),\boldsymbol{p}(t)), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0 \\
\frac{d\boldsymbol{p}}{dt} &= -G_x(\boldsymbol{x}(t),\boldsymbol{p}(t)), \quad \boldsymbol{p}(0) = \boldsymbol{p}_0 \\
\frac{dM}{dt} &= -(G_{xp})^T M - MG_{px} - MG_{pp}M - G_{xx}, \quad M(0) = i\epsilon I \\
\frac{da}{dt} &= -\frac{a}{2G}(c^2 trace(M) - G_x \cdot G_p - G_p^T M G_p, \quad A(0) = A_0
\end{aligned} \tag{4.29}$$

where the velocity term $c^2 = \lambda + 2\mu$ for the P-wave and $c^2 = \mu$ for the S-wave. The term G is the corresponding Hamiltonian. There is one extra equation about the direction $\boldsymbol{D}$ in the S-wave's dynamics,

$$\frac{dD_k}{dt} = -\left(\frac{dp_j(t)}{dt} D_j\right) \frac{p_k(t)}{|\boldsymbol{p}(t)|^2}, \quad \mathbf{D}(0) = \mathbf{D}_0.$$

The initial condition of the system above will be given by the Multiscale Gaussian Wavepacket transform, which will be specified in the later section. Now the way of propagating the phase functions $\tau^{P,S}$ and the amplitude vectors $\mathbf{A}^{P,S}$ is provided, and it allows us to finish the construction of a single-beam asymptotic solution,

$$\Phi^P(t, \boldsymbol{x}) = a(t) \nabla \tau(t, \boldsymbol{x}(t)) e^{i\omega\tau(t,\boldsymbol{x})} \tag{4.30}$$

$$\Phi^S(t, \boldsymbol{x}) = a(t) \boldsymbol{D} e^{i\omega\tau(t,\boldsymbol{x})}, \tag{4.31}$$

and the phase function is approximated by the Taylor expansion near the central ray,

$$\begin{aligned}\tau^{P,S}(t, \boldsymbol{x}) &= \nabla \tau^{P,S} \cdot (\boldsymbol{x} - \boldsymbol{x}(t)) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}(t))^T M^{P,S}(t)(\boldsymbol{x} - \boldsymbol{x}(t)) \\ &= \boldsymbol{p}(t) \cdot (\boldsymbol{x} - \boldsymbol{x}(t)) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}(t))^T M^{P,S}(t)(\boldsymbol{x} - \boldsymbol{x}(t)).\end{aligned} \tag{4.32}$$

The Gaussian profile is offered by the imaginary part of the Hessian matrix M

$$\exp\left(-\frac{\omega}{2}(\boldsymbol{x} - \boldsymbol{x}(t))^T Im(M(t))(\boldsymbol{x} - \boldsymbol{x}(t))\right). \tag{4.33}$$

Suggested by Lemma 4.2.1, a beam ansatz will be always well localized throughout the propagation.

4.3 Multiscale Gaussian Wavepacket Transform for Elastic Waves

The initial condition of the elastic wave equation (4.3) can be any general L_2 vector-valued function, and it is not necessary to take the exact form like,

$$\mathbf{A} \exp\left(i\omega\left(\boldsymbol{p}(0)^T(\boldsymbol{x}-\boldsymbol{x}_0)+\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_0)^T M(0)(\boldsymbol{x}-\boldsymbol{x}_0)\right)\right). \tag{4.34}$$

The problem here is that how to decompose any general L_2 function to multiple Gaussian wavepackets like the above form (4.34) efficiently and make the total number of beams to be calculated as small as possible.

We will provide a very brief introduction of the Multiscale Gaussian Wavepacket Transform [48] for the scalar functions first in this section. More details can be found in [48]. Then its extension designed for the vector-valued initial conditions is presented afterwards.

4.3.1 Multiscale Gaussian Wavepacket Transform: Vector Functions

After proposing the Multiscale Gaussian Wavepacket transform for the scalar function in Section 2.2, we would like to extend this idea to the vector-valued function $\boldsymbol{f}$. Here we assume that each component of the vector f_j is a L_2 function.

4.3.1.1 Decomposition of the Single Wavepacket

Suppose we have already applied the wavepacket transform to each component of the initial condition $\boldsymbol{f}$, i.e.

$$\boldsymbol{f} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_d \end{pmatrix}$$
$$= \sum_{l,i,k} \boldsymbol{\kappa}_{l,i,k} \phi_{l,i,k}.$$

The idea here is to decompose each single Gaussian wavepacket into the sum of the P-wave and the S-wave. Let the unit vector $\boldsymbol{v}_{l,i,k}$ be $\frac{\boldsymbol{\xi}_{l,i}}{|\boldsymbol{\xi}_{l,i}|}$, then

$$\boldsymbol{\kappa}_{l,i,k} = \left(\boldsymbol{\kappa}_{l,i,k}^T \boldsymbol{v}_{l,i,k}\right) \boldsymbol{v}_{l,i,k} + \left(I_d - \boldsymbol{v}_{l,i,k} (\boldsymbol{v}_{l,i,k})^T\right) \boldsymbol{\kappa}_{l,i,k}. \tag{4.35}$$

The first term on the right hand side of equation (4.35) is the P-wave component, and the initial condition for its amplitude vector can then be written as

$$\left(\boldsymbol{\kappa}_{l,i,k}^T \boldsymbol{v}_{l,i,k}\right) \boldsymbol{v}_{l,i,k} = \left(\boldsymbol{\kappa}_{l,i,k}^T \frac{\boldsymbol{\xi}_{l,i}}{|\boldsymbol{\xi}_{l,i}|^2}\right) \boldsymbol{\xi}_{l,i}. \tag{4.36}$$

To initialize the S-wave, we have to specify initial directions which are orthogonal to each other and all of them are supposed to be orthogonal to $\boldsymbol{v}_{l,i,k}$.

Therefore, we choose the first direction $\mathbf{D}^{(1)}$ as the unit vector of the first column vector of the matrix $\left(I_d - \boldsymbol{v}_{l,i,k}(\boldsymbol{v}_{l,i,k})^T\right)$ and apply the Gram-Schmidt process to generate the rest $\boldsymbol{D}^{(m)}$ for $m = 2, \cdots, d-1$. And we notice that every column vector of the matrix $\left(I_d - \boldsymbol{v}_{l,i,k}(\boldsymbol{v}_{l,i,k})^T\right)$ is orthogonal to $\boldsymbol{v}_{l,i,k}$, therefore, so are their linear combinations. The corresponding amplitude norm a_m,

$$a_m = \boldsymbol{\kappa}_{l,i,k}^T \left(I_d - \boldsymbol{v}_{l,i,k}(\boldsymbol{v}_{l,i,k})^T\right) \mathbf{D}^{(m)}, \quad m = 1, 2, \cdots, d-1. \tag{4.37}$$

4.3.1.2 Preprocessing the Initial Condition

Before applying the method described above, we pre-process the initial condition first. Following the same technology employed in [47], there are supposed to be two different branches $\boldsymbol{\kappa}_{l,i,k}^{+}$ and $\boldsymbol{\kappa}_{l,i,k}^{-}$ for each wavepacket corresponding to different signs of the Hamiltonian $\pm c(\boldsymbol{x})|\boldsymbol{p}|$, where $c(\boldsymbol{x})$ is the corresponding velocity.

To satisfy both the initial wavefield $\boldsymbol{f}$ and the initial velocity $\boldsymbol{g}$, we define

$$(\boldsymbol{\kappa}_{l,i,k}^{+} + \boldsymbol{\kappa}_{l,i,k}^{-})\phi_{l,i,k} = \boldsymbol{\kappa}_{l,i,k}\phi_{l,i,k} = \boldsymbol{f}. \tag{4.38}$$

Taking the derivative of the wavefield about the time variable t yields

$$(\boldsymbol{\kappa}_{l,i,k}^{-} - \boldsymbol{\kappa}_{l,i,k}^{+}) \left(i|\boldsymbol{\xi}_{l,i}|G^{+}(\boldsymbol{x}_0, \boldsymbol{p}_0)\right) \phi_{l,i,k} \approx \boldsymbol{\Xi}_{l,i,k}\phi_{l,i,k} = \mathbf{g}, \tag{4.39}$$

where $\boldsymbol{\Xi}_{l,i,k}$ is the coefficients generated from decomposing the initial velocity $\mathbf{g}$. Here the left hand side of equation (4.39) is not the complete form of the beam's time derivative, instead, we pick the leading order term to approximate.

After solving the coefficients $\boldsymbol{\kappa}^{+}_{l,i,k}$ and $\boldsymbol{\kappa}^{-}_{l,i,k}$ based on equation (4.38) and equation (4.39), we can apply the decomposition described by equations (4.35) - (4.37) to $\boldsymbol{\kappa}^{+}_{l,i,k}$ and $\boldsymbol{\kappa}^{-}_{l,i,k}$ respectively. The summary of the vector-version Multiscale Gaussian Wavepacket is provided below.

Algorithm 3 Discrete Vector-Valued Gaussian Wavepacket Transform

---

1.Call the Discrete Gaussian Wavepacket Transform for each component in the discrete signal **f** and **g**

2. Use equations (4.38) and (4.39) to compute $\boldsymbol{\kappa}^{+}_{l,i,k}$ and $\boldsymbol{\kappa}^{-}_{l,i,k}$
3. Generate P-wave with the amplitude vector $\left(\boldsymbol{\kappa}^{\pm}_{l,i,k} \cdot \frac{\boldsymbol{\xi}_{l,i}}{|\boldsymbol{\xi}_{l,i}|^2}\right)\boldsymbol{\xi}_{l,i}$
4. Generate S-waves $a_m\mathbf{D}^{(m)}$ by equation (4.37), for $m = 1, 2, \cdots, d-1$.

The above process defines the initial amplitude vectors for the P-wave and the S-wave, and the initial value of the phase function and its derivatives are given by the Multiscale Gaussian Wavepacket transform of the scalar form, that is

$$
\begin{aligned}
\frac{d\boldsymbol{x}}{dt} &= G_p(\boldsymbol{x}(t), \boldsymbol{p}(t)), \quad \boldsymbol{x}(0) = \frac{k}{L_l} \\
\frac{d\boldsymbol{p}}{dt} &= -G_x(\boldsymbol{x}(t), \boldsymbol{p}(t)), \quad \boldsymbol{p}(0) = 2\pi\frac{\xi_{l,i}}{|\xi_{l,i}|}, \\
\frac{dM}{dt} &= -(G_{xp})^T M - MG_{px} - MG_{pp}M - G_{xx}, \quad M(0) = i(2\pi^2\sigma_l^2/|\xi_{l,i}|)I. \\
\frac{da}{dt} &= -\frac{a}{2G}(c^2 trace(M) - G_x \cdot G_p - G_p^T M G_p, \quad a(0) = \left(\sqrt{\frac{\pi}{L_l N}}\sigma_l\right)^d. \\
\frac{dD_j}{dt} &= (\frac{d\tau_{,k}}{dt}D_k)\frac{\tau_{,j}}{|\nabla\tau|^2} \quad \boldsymbol{D}^{(m)}(0) = \boldsymbol{D}^{(m)}.
\end{aligned}
$$

4.4 Multiscale Gaussian Beam Method for Periodic Boundary Value Problem

In the above section, we have demonstrated the way to decompose the vector-valued initial conditions in the L_2 space. To solve the periodic boundary problem,each parameters and solutions are assumed to be periodic functions. Meanwhile, the central ray in the periodic boundary problem will be smooth along the propagation in the sense of modules.

The principle shown in Figure 4.1 will be employed to solve the periodic boundary value problem. The red dashed line represents the wavepacket leaving the domain $[0, 0.5]$, as

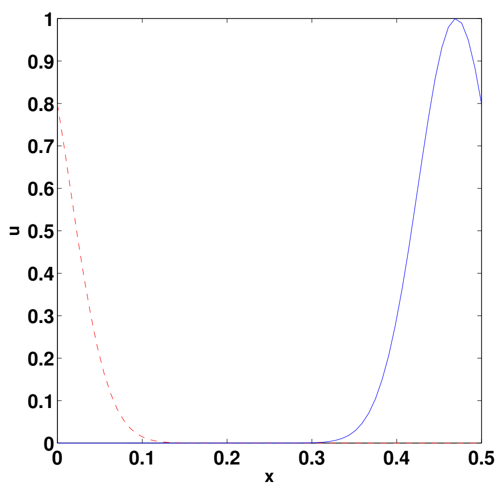


Figure 4.1: Periodic Boundary Problem: The Case Wavepacket leaving the Boundary.

the left half goes beyond the domain, while the right half still shows up. The missing left half of the beam solution will enter from the other side with the same shape, which is the blue line in the graph suggested by the periodic boundary condition. The cubic region $[0, 0.5]^3$ is chosen to test the correctness of our algorithm. We show numerical results in Section 4.8.

4.5 Multiscale Gaussian beam method for Homogeneous Dirichlet Boundary Condition

In this section, we would like to explore the solution to the homogeneous Dirichlet Boundary condition. From now on, we use the 3-D space $\{\boldsymbol{x} = (x, y, z) : x, y, z \in \mathbb{R}\}$ as our model.

The wavefield on the boundary is assumed to be zero in this section. When the reflection happens, the sum of all wavefields at time t^r and the central point $\boldsymbol{x}(t^r)$ of the ray should vanish, i.e. $\boldsymbol{u}(t^r, \boldsymbol{x}(t^r)) = 0$. The time when the central point $\boldsymbol{x}(t^r)$ of the ray is on the boundary is defined as the reflection time. From now on, all equations below in this section are defined on the point $(t^r, \boldsymbol{x}(t^r))$, if not specified. The Hamiltonian used in this section is assumed to be positive $G = c(\boldsymbol{x})|\boldsymbol{p}|$, and the negative Hamiltonian will be treated similarly.

4.5.1 P-wave Reflecting Beams: Ray Direction

When the P-wave reflection happens, the total wavefield is made up by three different sources, the original P-wave Gaussian beam, the new P-wave beam after the reflection (PP-wave) and the new S-wave beam (PS-wave). At the reflection point $\boldsymbol{x}(t^r)$,

$$-a^P e^{i\tau^P} \nabla \tau^P = a^{PP} e^{i\tau^{PP}} \nabla \tau^{PP} + a^{PS} e^{i\tau^{PS}} \mathbf{D}^{PS}. \tag{4.40}$$

Both P and S-wave will be generated after reflection.

All phase functions should have the same value at the reflection point, $\tau^P = \tau^{PP} =$

τ^{PS}. Otherwise, if we change the value of large wavenumber $|\boldsymbol{\xi}_{l,i}|$, the homogeneous boundary condition will be violated.

The principle to derive new phase functions τ^{PP} and τ^{PS} is to take advantage of the continuity conditions, i.e. the continuity of the tangential components of the first order derivatives of τ, so

$$\begin{aligned} \tau_y^P &= \tau_y^{PP} = \tau_y^{PS}, \\ \tau_z^P &= \tau_z^{PP} = \tau_z^{PS}, \end{aligned} \tag{4.41}$$

where we assume the reflection happens along the surface $\{\boldsymbol{x} = (x, y, z) : x = 0\}$. Besides the spatial variables, the partial derivative of the phase function with respect to the time variable t should also follow,

$$\begin{aligned} \dot{\tau}^P &= \dot{\tau}^{PP} = \dot{\tau}^{PS} \\ \Rightarrow c^P|\nabla\tau^P| &= c^S|\nabla\tau^{PS}|, \\ c^P|\nabla\tau^P| &= c^P|\nabla\tau^{PP}|. \end{aligned} \tag{4.42}$$

where $c^P = \sqrt{\lambda + 2\mu}$ is the velocity of the P-wave and $c^S = \sqrt{\mu}$ is the velocity of the S-wave. The partial derivatives along the tangential directions of the boundary can be obtained directly from equation (4.41). To obtain the momentum along the reflection direction or the normal direction of the boundary, one needs to use equations (4.42) and (4.41) collectively.

$$\tau_x^{PS} = -sign(\tau_x^P)\sqrt{\left(\frac{c^P}{c^S}\right)^2 |\nabla\tau^P|^2 - (\tau_y^{PS})^2 - (\tau_z^{PS})^2}. \tag{4.43}$$

Here we focus our derivation on the reflection from P-wave to S-wave, while the rule of the reflection between the same mode can follow the same way, so we skip it here.

The only ambiguity left here is the case when a beam hits the boundary at a corner since it causes diffraction and the above derivation does not apply any more. Here we simply ignore the situation when a beam hits a corner of the domain since the Gaussian method is asymptotic. The numerical accuracy will not be degraded without those beams as those diffractions have exponentially small effects.

4.5.2 P-wave Reflecting Beams: The Hessian of the Phase

To illustrate the derivation of the second order derivative terms, we pick three entries among six distinct entries in the Hessian for explanation, τ_{yy}, τ_{xy} and τ_{xx}, since all other entries can be classified into one of these three types. Again, the reflection is assumed to happen along the surface $\{\boldsymbol{x} = (x, y, z) : x = 0\}$ and all terms without arguments are defined at the reflection point.

To start with the first type τ_{yy}, which is tangential component

$$\tau_{yy}^{P} = \tau_{yy}^{PS} = \tau_{yy}^{PP}. \tag{4.44}$$

τ_{zz} and τ_{yz} will also stay the same.

To derive the second type of terms τ_{xy}^{PS}, we use the partial derivatives about the time variable t,

$$\dot{\tau}_y^P = c_y^P|\nabla\tau^P| + c^P\frac{(\nabla\tau_y^P)^T\nabla\tau^P}{|\nabla\tau^P|} \tag{4.45}$$

$$\dot{\tau}_y^{PS} = c_y^S|\nabla\tau^{PS}| + c^S\frac{(\nabla\tau_y^{PS})^T\nabla\tau^{PS}}{|\nabla\tau^{PS}|}, \tag{4.46}$$

The notation used here is $\dot{\tau}_y = \frac{\partial^2\tau}{\partial t\partial y}$. We have

$$\dot{\tau}_y^P = \dot{\tau}_y^{PS}, \tag{4.47}$$

due to the continuity of the tangential and time component. Now substitute equation (4.45) and equation (4.46) into equation (4.47),

$$\begin{aligned}
c_y^P|\nabla\tau^P| + c^P\frac{(\nabla\tau_y^P)^T\nabla\tau^P}{|\nabla\tau^P|} &= c_y^S|\nabla\tau^{PS}| + c^S\frac{(\nabla\tau_y^{PS})^T\nabla\tau^{PS}}{|\nabla\tau^{PS}|} \\
(\nabla\tau_y^{PS})^T\nabla\tau^{PS} &= \frac{|\nabla\tau^{PS}|}{c^S}\left(c_y^P|\nabla\tau^P| + c^P\frac{(\nabla\tau_y^P)^T\nabla\tau^P}{|\nabla\tau^P|} - c_y^S|\nabla\tau^{PS}|\right) \\
\tau_{xy}^{PS} &= \frac{|\nabla\tau^{PS}|}{c^S\tau_x^{PS}}\left(c_y^P|\nabla\tau^P| + c^P\frac{\nabla\tau_y^P\cdot\nabla\tau^P}{|\nabla\tau^P|} - c_y^S|\nabla\tau^{PS}|\right) \\
&\quad - \frac{1}{\tau_x^{PS}}(\tau_y^{PS}\tau_{yy}^{PS} + \tau_z^{PS}\tau_{yz}^{PS}).
\end{aligned} \tag{4.48}$$

To obtain the last type of the term τ_{xx}^{PS}, we need to derive the formula of τ_{tx}^{PS} first.

$$\dot{\tau}^P = c^P|\nabla\tau^P| \Rightarrow \ddot{\tau}^P = c^P\frac{(\nabla\dot{\tau}^P)^T\nabla\tau^P}{|\nabla\tau^P|}, \tag{4.49}$$

$$\dot{\tau}^{PS} = c^S|\nabla\tau^{PS}| \Rightarrow \ddot{\tau}^{PS} = c^S\frac{(\nabla\dot{\tau}^{PS})^T\nabla\tau^{PS}}{|\nabla\tau^{PS}|}. \tag{4.50}$$

We then have the following equation from equation (4.50) and equation (4.49),

$$
\begin{aligned}
c^P \frac{(\nabla \dot{\tau}^P)^T \nabla \tau^P}{|\nabla \tau^P|} &= c^S \frac{(\nabla \dot{\tau}^{PS})^T \nabla \tau^{PS}}{|\nabla \tau^{PS}|} \\
(\nabla \dot{\tau}^{PS})^T \nabla \tau^{PS} &= \frac{|\nabla \tau^{PS}|}{c^S} c^P \frac{(\nabla \dot{\tau}^P)^T \nabla \tau^P}{|\nabla \tau^P|} \\
\dot{\tau}_x^{PS} &= \frac{1}{\tau_x^{PS}} \left(\frac{|\nabla \tau^{PS}|}{c^S} c^P \frac{(\nabla \dot{\tau}^P)^T \nabla \tau^P}{|\nabla \tau^P|} - \dot{\tau}_y^{PS} \tau_y^{PS} - \dot{\tau}_z^{PS} \tau_z^{PS} \right) \qquad (4.51)
\end{aligned}
$$

With the formula about the term $\dot{\tau}_x^{PS}$ given above, the term τ_{xx}^{PS} can be obtained by solving the following equation,

$$
\dot{\tau}^{PS} = c^S |\nabla \tau^{PS}| \Rightarrow \dot{\tau}_x^{PS} = c_x^S |\nabla \tau^{PS}| + c^S \frac{(\nabla \tau_x^{PS})^T \nabla \tau^{PS}}{|\nabla \tau^{PS}|}. \qquad (4.52)
$$

To remark, τ_x^{PS} will not be zero as we assumed our initial conditions are compactly supported.

4.5.3 P-wave Reflecting Beams: Amplitude Vector

We have so far already derived the initial condition of all terms involved with the phase function for our new ODE dynamic system after P-wave reflection. Since the phase function itself does not change after reflection in the center of the beam, we will have the following equation about the amplitude to satisfy the homogeneous boundary condition at the reflection point $\boldsymbol{x}(t^r)$,

$$
-a^P \nabla \tau^P = a^{PP} \nabla \tau^{PP} + a^{PS} \mathbf{D}^{PS}. \qquad (4.53)
$$

We know that the PS-wave's amplitude direction $\mathbf{D}^{PS}$ is orthogonal to its ray direction $\nabla\tau^{PS}$ by the definition and with unit norm by our restriction. Consequently, the following equation can be obtained if we project both sides of equation (4.53) to the vector $\nabla\tau^{PS}$ at the same time.

$$\begin{aligned} -a^P(\nabla\tau^P)^T\nabla\tau^{PS} &= a^{PP}(\nabla\tau^{PP})^T\nabla\tau^{PS}, \\ a^{PP} &= -a^P\frac{(\nabla\tau^P)^T\nabla\tau^{PS}}{(\nabla\tau^{PP})^T\nabla\tau^{PS}}. \end{aligned} \tag{4.54}$$

Like the initial condition, $a^{PS}\mathbf{D}^{PS}$ is the summation of two S-waves.

$$a^{PS}\mathbf{D}^{PS} = \sum_i \alpha_i^{PS}\boldsymbol{D}^{(i)}. \tag{4.55}$$

Similar to the initial condition, we pick the first direction $\boldsymbol{D}^{(1)}$ to be the first column vector of the matrix $I_3 - \boldsymbol{v}\boldsymbol{v}^T$ and $\boldsymbol{v} = \frac{\nabla\tau^{PS}}{|\nabla\tau^{PS}|}$. Then, the second vector $\boldsymbol{D}_2$ will be $\boldsymbol{D}^{(2)} = \boldsymbol{D}^{(1)} \times \boldsymbol{v}$, where $\times$ represents the cross product between two vectors. After normalizing each direction, we can project the residual $-a^P\nabla\tau^P - a^{PP}\nabla\tau^{PP}$ to each direction to obtain the amplitude α_i^{PS}.

4.5.4 S-wave Reflecting Beams: the Phase term

Similar to the P-wave reflection, we have the following equation to satisfy the homogeneous boundary condition for the S-wave wavepacket at the reflection point $\boldsymbol{x}(t^r)$.

$$-a^S e^{i\tau^S}\mathbf{D}^S = a^{SP}e^{i\tau^{SP}}\nabla\tau^{SP} + a^{SS}e^{i\tau^{SS}}\mathbf{D}^{SS}. \tag{4.56}$$

After taking first glance of the above equation, the S-wave reflection dynamics seems to be the same as the P-wave reflection developed in the last section. However, we will find that there will be some significant difference between the S-wave reflection and the P-wave reflection. To begin with, let us still use the case hitting the surface $\{\boldsymbol{x} = (x, y, z) : x = 0\}$ as an example,

$$\tau_y^S = \tau_y^{SS} = \tau_y^{SP}, \tag{4.57}$$

$$\tau_z^S = \tau_z^{SS} = \tau_z^{SP}, \tag{4.58}$$

$$\tau_t^S = \tau_t^{SS} = \tau_t^{SP}, \tag{4.59}$$

$$\Rightarrow c^S|\nabla\tau^S| = c^P|\nabla\tau^{SP}|. \tag{4.60}$$

With all these equations combined, we will get

$$\tau_x^{SP} = -sign(\tau_x^S)\sqrt{\left(\frac{c^S|\nabla\tau^S|}{c^P}\right)^2 - (\tau_y^S)^2 - (\tau_z^S)^2}, \tag{4.61}$$

and the S-wave's velocity $c^S = \sqrt{\mu}$ is less than the P-wave's $c^P = \sqrt{\lambda + 2\mu}$. This leads to the possibility that the part inside the square root in equation (4.61) will be negative, or equivalently, τ_x^{SP} can be in general a pure imaginary number. It will be a disaster, since there will be some exponentially increasing wave on one side of the boundary.

Let's consider the regular situation first, in which case, equation (4.61) is a real number. It is same as the one employed in the P-wave reflection as illustrated above, so we skip it here.

The second case is when the term τ_x^{SP} is a pure imaginary number, i.e.

$$\nabla\tau^{sp} = \begin{pmatrix} i\sqrt{-\left(\frac{c^S|\nabla\tau^S|}{c^P}\right)^2 + (\tau_y^S)^2 + (\tau_z^S)^2} \\ \tau_y^S \\ \tau_z^S \end{pmatrix}. \tag{4.62}$$

This phenomenon is called the evanescent wave and the energy will fade away quickly around the boundary in this case. Therefore, there is no need to derive its Hessian due to its small energy.

4.5.5 S-wave Reflecting Beams: the Amplitude Vector

The evanescent wave fades away quickly, however, we will still include SP-wave's amplitude vector in our derivation so as to make the derivation easier. Moreover, we need the nonzero amplitude vector $\boldsymbol{A}^{SP}$ to make the homogeneous boundary assumption true at the reflection point $\boldsymbol{x}(t^r)$. To summarize,

$$-a^S\mathbf{D}^S = a^{SP}(Re(\nabla\tau^{sp}) + iIm(\nabla\tau^{sp})) + a^{SS}\mathbf{D}^{SS}. \tag{4.63}$$

The way we get the amplitude vector $\boldsymbol{A}^{SP}$ is still from the same idea by using the fact that the SS wave's amplitude direction $\mathbf{D}^{SS}$ is orthogonal to its ray direction $\mathbf{v}$,

$$\left(Re(\boldsymbol{A}^{SP})Re(\nabla\tau^{SP}) - Im(\boldsymbol{A}^{SP})Im(\nabla\tau^{SP})\right)\cdot\boldsymbol{v} = -Re(a^S)\mathbf{D}^S\cdot\boldsymbol{v}; \tag{4.64}$$

$$\left(Re(\boldsymbol{A}^{SP})Im(\nabla\tau^{SP}) + Im(\boldsymbol{A}^{SP})Re(\nabla\tau^{SP})\right)\cdot\boldsymbol{v} = -Im(a^S)\mathbf{D}^S\cdot\boldsymbol{v}. \tag{4.65}$$

The amplitude of the SP-wave's amplitude $\boldsymbol{A}^{SP}$ can be obtained by solving the above system. Consequently, the residual $-a^S D^S - a^{SP}\nabla\tau^{SP}$ is now well defined. Following the same process defined in the PS wave case, we can set up the amplitudes and directions easily for SS-wave.

Notice that the ray after reflection is no longer smooth, which means that Lemma 4.2.1 is not applicable when reflections happen. Naturally, one needs to show that the imaginary part of the Hessian after reflections defined above will still be positive definite, especially for the PS wave and the SP wave case. In [4], authors have already proved this is true for the PP wave and the SS wave, i.e. the conversion between same wave modes. The proof about the conversion between different wave modes is provided in Appendix B.

4.5.6 Method of Images for Boundary Conditions

In [4], authors have proposed a method to tackle the problem caused by partially reflected beams. The partial reflection problem means the frontier part of a beam is needed to be reflected back even when its central ray has not hit the boundary yet and consequently the reflection dynamics has not been called. This is due to the fact that beams have nonzero width and illustrated in the following graph.

Therefore, some modifications should be added to these partial reflection cases so that the homogeneous boundary condition is always satisfied as well as our wavefield remains to be continuous. Our strategy presented here is that the outer part is considered to be reflected back, which is carried by some artificial beams. So we essentially apply the odd extension to those beams, like what Figure 4.3 shows. The trajectory of the

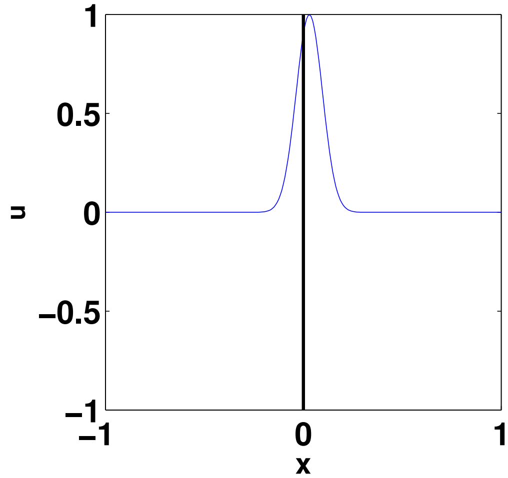


Figure 4.2: Partially Reflected beams

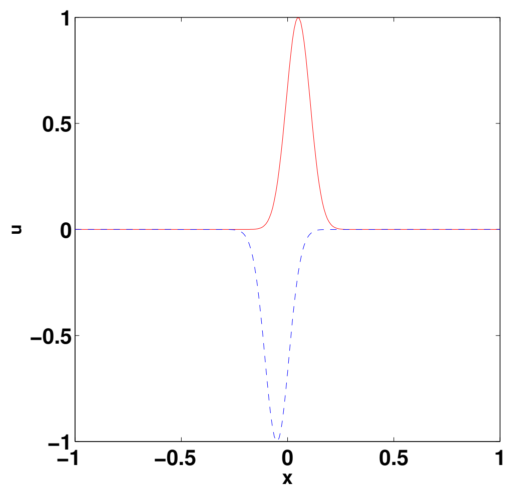


Figure 4.3: Partially Reflected Beams with Odd Extension

blue dashed wavepacket in the graph is completely determined by its associates, the red solid wavepacket in the graph. It implies that we don't use any extra assumption of the velocity outside the domain. The blue dashed wavepacket only serves as a supplementary beam to satisfy the vanishing boundary condition.

4.6 Stationary Phase Analysis of Beams

To reinitialize or sharpen the single Gaussian beam ansatz, we need to apply some stationary phase analysis to the single beam. So we first list some lemmas and computations here which are needed to implement the reinitialization process in the next section.

For any function u in $L^2(\mathbb{R}^d)$, there is a phase space decomposition method such that,

$$u(\boldsymbol{x}) = \left(\frac{\omega}{2\pi}\right)^{3d/2} \int_{R^{3d}} 2^{d/2} e^{i\omega(\boldsymbol{p}(\boldsymbol{x}-\boldsymbol{x}')-\boldsymbol{p}(\tilde{\boldsymbol{x}}-\boldsymbol{x}'))} e^{-\frac{\omega|\boldsymbol{x}-\boldsymbol{x}'|^2}{2}} e^{-\frac{\omega|\tilde{\boldsymbol{x}}-\boldsymbol{x}'|^2}{2}} u(\tilde{\boldsymbol{x}}) d\tilde{\boldsymbol{x}} d\boldsymbol{p} d\boldsymbol{x}'. \tag{4.66}$$

Here $\boldsymbol{x}$, $\tilde{\boldsymbol{x}}$ and $\boldsymbol{x}'$ are points in the spatial space $\mathbb{R}^d$ and $\boldsymbol{p}$ is the dual momentum variable in the frequency space. ω is a fixed parameter determining the size of Gaussian window functions.Here we would like to explore how to apply representation (4.66) to Gaussian beams without considering its amplitude's direction. Using P-wave as an example, consider $\Phi^P(t,\boldsymbol{x}) = a(t)e^{i\tau(t,\boldsymbol{x})}$ instead of $a(t)e^{i\tau(t,\boldsymbol{x})}\mathbf{p}(t)$. Moreover, all beam functions considered here are treated as single-variable functions by assuming a principal variable while other variables are fixed.

In all the following derivations, the principal variable is assumed to be the variable y, while other variables x and z are fixed.

4.6.1 Stationary Phase Approximation with Respect to Spatial Variables

The Gaussian beams u with y as the principal variable considered in this section are in the following form,

$$u(y;x,z,t) = A(x,z,t)e^{i\Phi(y-y_0)}e^{-\frac{1}{2}Im(\tau_{yy})(y-y_0^*)^2}, \tag{4.67}$$

where $\big(x_0, y_0, z_0\big)$ is the central point of the beam ansatz. All functions below are defined at this point if not specified, and

$$\Phi(y-y_0) = \tau_y(y-y_0)+\frac{Re(\tau_{yy})(y-y_0)^2}{2}+Re(\tau_{xy})(y-y_0)(x-x_0)+Re(\tau_{zy})(y-y_0)(z-z_0). \tag{4.68}$$

τ is the phase function. To define the scalar y_0^*, we have the following equation where all second order derivative terms are imaginary part only,

$$\begin{aligned}
&\Big(x-x_0, y-y_0, z-z_0\Big)\begin{pmatrix}\tau_{xx} & \tau_{xy} & \tau_{xz}\\ \tau_{xy} & \tau_{yy} & \tau_{yz}\\ \tau_{xz} & \tau_{zy} & \tau_{zz}\end{pmatrix}\begin{pmatrix}x-x_0\\ y-y_0\\ z-z_0\end{pmatrix}\\
&= \tau_{yy}(y-y_0)^2 + 2(\tau_{xy}(x-x_0)+\tau_{xz}(z-z_0))(y-y_0) + B(x,z)\\
&= \tau_{yy}\left(y-y_0+\frac{\tau_{xy}(x-x_0)+\tau_{xz}(z-z_0)}{\tau_{yy}}\right)^2 + \tilde{B}(x,z).
\end{aligned} \tag{4.69}$$

The complete square term in equation (4.69) is defined as $(y-y_0^*)^2$ and $\tilde{B}(x,z)$ is a constant with respect to y.

To make the expression clear, we denote $R = Re(\tau_{yy})$ and $I = Im(\tau_{yy})$ throughout this section. Applying decomposition (4.66) to $u(y; x, z, t)$, we have

$$u(y; x, z, t) = A(x, z, t)\left(\frac{\omega}{2\pi}\right)^{3/2} \int_{R^2} \sqrt{2} e^{i\omega p(y-y^{'})} e^{-\frac{\omega}{2}|y-y^{'}|^2} \psi(p, y^{'}) dp dy^{'}, \quad (4.70)$$

and

$$\psi(p, y^{'}) = e^{i\tilde{B}(x,z)} \int e^{-i\omega p(\tilde{y}-y^{'})} e^{-\frac{\omega}{2}|\tilde{y}-y^{'}|^2} e^{i\Phi(\tilde{y}-y_0)} e^{-\frac{I}{2}|\tilde{y}-y_0^*|^2} d\tilde{y}. \quad (4.71)$$

We will apply the stationary phase approximation described in the following lemma to calculate $\psi(p, y^{'})$.

Lemma 4.6.1. *We consider the behavior of* $I(\omega) = \int_a^b f(t) e^{i\omega g(t)} dt$, *where* f *and* g *are smooth enough to admit Taylor approximation near some appropriate points in* $[a, b]$, *and* g *is real-valued. Suppose there is some point* $c_0 \in [a, b]$ *and* $g^{'}(t) \neq 0$ *everywhere else in the closed interval* $[a, b]$. *Moreover,* $g^{''}(c_0) \neq 0$. *When* $\omega >> 1$, *we have*

$$I(\omega) = \int_a^b f(t) e^{i\omega g(t)} dt = f(c_0) e^{i\omega g(c_0)} e^{i\pi\delta/4} \sqrt{\frac{2\pi}{\omega |g^{''}(c_0)|}} + O\left(\frac{1}{\omega}\right), \quad (4.72)$$

where δ *is the sign of* $g^{''}(c_0)$.

To use Lemma 4.6.1, we first substitute $y^{'} = y_0 + ma$ into $u(y; x, z, t)$, and the value of a and ω will be defined later. By equation (4.71), we have

$$e^{-i\tilde{B}(x,z)} \psi(p, y_0 + ma) = \int e^{i\omega\left(\frac{\Phi(\tilde{y}-y_0)}{\omega} - p(\tilde{y}-y_0-ma)\right)} e^{-\frac{\omega}{2}|\tilde{y}-y_0-ma|^2} e^{-\frac{I}{2}|\tilde{y}-y_0^*|^2} d\tilde{y}.$$
(4.73)

To compute the critical point $\tilde{y}_0$ of the phase term $\frac{\Phi(\tilde{y}-y_0)}{\omega} - p(\tilde{y} - y_0 - ma)$, we have

$$
\begin{aligned}
0 &= \frac{\Phi'(\tilde{y} - y_0)}{\omega} - p, \\
0 &= \frac{R(\tilde{y} - y_0) + Re(\tau_{xy})(x - x_0) + Re(\tau_{yz})(z - z_0)}{\omega} + \frac{\tau_y}{\omega} - p. \qquad (4.74)
\end{aligned}
$$

that is

$$
\tilde{y}_0 = y_0 + \frac{\omega p - \tau_y - Re(\tau_{xy})(x - x_0) - Re(\tau_{yz})(z - z_0)}{R}. \qquad (4.75)
$$

We denote

$$
\begin{aligned}
E(x, z) &= \frac{-Re(\tau_{xy})(x - x_0) - Re(\tau_{yz})(z - z_0)}{R}, \qquad (4.76) \\
\tilde{y}_0 - y_0 &= \frac{\omega p - \tau_y}{R} + E.
\end{aligned}
$$

Notice that $E(x, z)$ is independent of the variable y, together with Lemma 4.6.1, the term $\psi(p, y_0 + ma)$ is equal to

$$
\begin{aligned}
e^{-i\tilde{B}(x,z)}\psi(p, y_0 + ma) = {} & e^{i\omega\left(\frac{\Phi(\frac{\omega p - \tau_y}{R} + E)}{\omega} - p(\frac{\omega p - \tau_y}{R} + E - ma)\right)} e^{-\frac{\omega}{2}|\frac{\omega p - \tau_y}{R} + E - ma|^2} \\
& \times e^{-\frac{I}{2}|\frac{\omega p - \tau_y}{R} + E - y_0^* + y_0|^2} e^{i\frac{\pi}{4}\frac{R}{|R|}} \sqrt{\frac{2\pi}{|R|}}.
\end{aligned}
$$

Substitute the expression of the term Φ into the above equation,

$$
\begin{aligned}
\psi(p, y_0 + ma) = {} & e^{i\tilde{B}} e^{i\frac{\pi}{4}\frac{R}{|R|}} \sqrt{\frac{2\pi}{|R|}} e^{-\frac{I}{2}|\frac{\omega p - \tau_y}{R} - y_0^* + y_0 + E|^2} e^{-\frac{\omega}{2}|\frac{\omega p - \tau_y}{R} + E - ma|^2} \\
& \times e^{i\omega\left(\frac{\tau_y}{\omega}\left(\frac{\omega p - \tau_y}{R} + E\right) - (\frac{R \cdot E}{\omega})\left(\frac{\omega p - \tau_y}{R} + E\right) - p\left(\frac{\omega p - \tau_y}{R} + E - ma\right) + \frac{R}{2\omega}(\frac{\omega p - \tau_y}{R} + E)^2\right)}. \qquad (4.77)
\end{aligned}
$$

We insert equation (4.77) into equation (4.70),

$$u(y;x,z,t) = aA(x,z,t)\sqrt{\left(\frac{\omega}{2\pi}\right)^3}\int \sqrt{2}e^{i\omega p(y-y_0-ma)}e^{-\frac{\omega}{2}|y-y_0-ma|^2}\psi(p,y_0+ma)dpdm$$
$$\approx 2Aa\sqrt{\left(\frac{\omega}{2\pi}\right)^3}\sqrt{\frac{\pi}{|R|}}e^{i\frac{\pi}{4}\frac{R}{|R|}}e^{i\tilde{B}}\int e^{i\omega p(y-y_0-ma)}e^{-\frac{\omega}{2}|y-y_0-ma|^2}e^{-\frac{I}{2}|\frac{\omega p-\tau y}{R}-y_0^*+y_0+E|^2}$$
$$\times\, e^{-\frac{\omega}{2}|\frac{\omega p-\tau y}{R}+E-ma|^2}e^{i\omega\left(\frac{\tau y}{\omega}\left(\frac{\omega p-\tau y}{R}+E\right)-\frac{R\cdot E}{\omega}\left(\frac{\omega p-\tau y}{R}+E\right)-p\left(\frac{\omega p-\tau y}{R}+E-ma\right)\right)}$$
$$\times\, e^{i\frac{R}{2}\left(\frac{\omega p-\tau y}{R}+E\right)^2}dpdm. \tag{4.78}$$

By applying the stationary phase approximation to the variable $\tilde{y}$, we reduce the triple integral (4.66) to the double integral (4.78).

Notice that in Lemma 4.6.1, the assumption $R \neq 0$ is required. In general, it is not true. However, if $R = 0$, then the integral about the function $\psi(p, y')$ is nothing but the Fourier transform of a Gaussian function about y, so we omit it here.

4.6.2 Stationary Phase Approximation with Respect to Momentum Variables

Starting from the double integral (4.78), we would like to apply the stationary phase approximation again. However, this time the variable we apply over is the momentum variable p.

$$u(y;x,z,t) = A(x,z,t)e^{i\tilde{B}}\sqrt{\left(\frac{\omega}{2\pi}\right)^3}\sqrt{\frac{4\pi}{|R|}}e^{i\frac{\pi}{4}\frac{R}{|R|}}a\int e^{-\frac{\omega}{2}|y-y_0-ma|^2}dm$$
$$\times\int e^{i\omega p(y-y_0-ma)}e^{-\frac{I}{2}|\frac{\omega p-\tau y}{R}-y_0^*+y_0+E|^2}e^{-\frac{\omega}{2}|\frac{\omega p-\tau y}{R}+E-ma|^2}$$
$$\times e^{i\omega\left(\frac{\tau y}{\omega}\left(\frac{\omega p-\tau y}{R}+E\right)-\frac{R\cdot E}{\omega}\left(\frac{\omega p-\tau y}{R}+E\right)-p\left(\frac{\omega p-\tau y}{R}+E-ma\right)+\frac{R}{2\omega}\left(\frac{\omega p-\tau y}{R}+E\right)^2\right)}dp. \tag{4.79}$$

For the inner integral of equation (4.79), the phase function $g(p)$ is

$$g(p) = p(y-y_0-ma)+\frac{\tau y}{\omega}\left(\frac{\omega p-\tau y}{R}+E\right)$$
$$-\frac{R\cdot E}{\omega}\left(E+\frac{\omega p-\tau y}{R}\right)-p\left(E-ma+\frac{\omega p-\tau y}{R}\right)+\frac{R}{2\omega}\left(\frac{\omega p-\tau y}{R}+E\right)^2,$$

and the smooth function $f(p)$ is,

$$f(p) = e^{-\frac{\omega}{2}|\frac{\omega p-\tau y}{R}+E-ma|^2}e^{-\frac{I}{2}|\frac{\omega p-\tau y}{R}-y_0^*+y_0+E|^2}. \tag{4.80}$$

To compute the critical point,

$$g'(p) = y-y_0-ma+\frac{\tau y}{R}-E-\left(E-ma+\frac{\omega p-\tau y}{R}\right)-\frac{p\omega}{R}+\left(\frac{\omega p-\tau y}{R}+E\right), \tag{4.81}$$

then the critical point p_0 is

$$
\begin{aligned}
\frac{\omega p}{R} &= y - y_0 - E + \frac{\tau y}{R} \\
\Rightarrow p_0 &= \frac{R}{\omega}\big(y - y_0 - E\big) + \frac{\tau y}{\omega} \\
\Rightarrow \frac{\omega p_0 - \tau y}{R} &= y - y_0 - E. \qquad (4.82)
\end{aligned}
$$

The second order derivative of the phase function g at the critical point p_0 is

$$
g''(p_0) = -\frac{\omega}{R}. \qquad (4.83)
$$

So the approximation of the inner integral of equation (4.79) is

$$
\begin{aligned}
&\int e^{i\omega p(y-y_0-ma)} e^{-\frac{I}{2}|\frac{\omega p-\tau y}{R}-y_0^*+y_0+E|^2} e^{-\frac{\omega}{2}|\frac{\omega p-\tau y}{R}+E-ma|^2} \\
&\times e^{i\omega\left(\frac{\tau y}{\omega}\left(\frac{\omega p-\tau y}{R}+E\right)-\frac{R\cdot E}{\omega}\left(\frac{\omega p-\tau y}{R}+E\right)-p\left(\frac{\omega p-\tau y}{R}+E-ma\right)+\frac{R}{2\omega}\left(\frac{\omega p-\tau y}{R}+E\right)^2\right)} dp \\
&= \\
&\sqrt{\frac{\pi R}{\omega^2}} e^{-i\frac{\pi R}{4|R|}} e^{i\omega p_0(y-y_0-ma)} e^{-\frac{I}{2}|y-y_0^*|^2} e^{-\frac{\omega}{2}|y-y_0-ma|^2} \\
&\times e^{i\omega\left((y-y_0)\left(\frac{\tau y-R\cdot E}{\omega}\right)\right)} e^{i\omega\left(-p_0(y-y_0-ma)+\frac{R}{2\omega}(y-y_0)^2\right)}.
\end{aligned}
$$

We summarize the stationary analysis conducted in this section in the following lemma,

Lemma 4.6.2. *Suppose the principal variable selected for the function $u(y;x,z,t)$ is y*

$$
u(y;x,z,t) = A(x,z,t)e^{i\Phi(y-y_0)}e^{-\frac{1}{2}Im(\tau yy)(y-y_0^*)^2}, \qquad (4.84)
$$

and $R = Re(\tau_{yy})$ is nonzero, then the following decomposition can be obtained,

$$u(y; x, z, t) = aA(x, z, t)e^{i\tilde{B}(x,z)}\sqrt{\frac{\omega}{2\pi}}$$
$$\int e^{-\omega|y-y_0-ma|^2} e^{-\frac{I}{2}|y-y_0^*|^2} e^{i\tilde{\tau}_y(y-y_0)} e^{i\frac{R}{2}(y-y_0)^2} dm + O\left(\frac{1}{\omega}\right) \quad (4.85)$$

where $\tilde{\tau}_y = \tau_y - R \cdot E$ and

$$R \cdot E = -Re(\tau_{xy})(x - x_0) - Re(\tau_{yz})(z - z_0). \quad (4.86)$$

The fixed parameter a is defined as $a = \frac{1}{\sqrt{I}}$. The value of ω will be specified later, but its order is $O(I)$.

4.7 Sharpening Beams by Reinitialization

In this section, we would like to propose a new reinitialization strategy based on Lemma 4.6.2 from Section 4.6. Again, we base our proof on the assumption that the variable y is the principal variable. We first illustrate the reason why proposing a new reinitialization strategy is necessary.

4.7.1 The First Motivation for Developing a New Reinitialization Strategy

It is necessary to add a reinitialization process into the propagation since the width of a beam will increase exponentially in some generic medium [48]. We use the linear velocity and 1-D problem for explanation. Suppose $c(x) = \alpha + \beta x$ where α and β are constants, then the Riccati equation about the Hessian M will be

$$\frac{dM}{dt} + 2M\beta = 0, \quad M(0) = i\epsilon. \tag{4.87}$$

Solving this simple linear ODE, we have

$$M(t) = i\epsilon e^{-2\beta t}. \tag{4.88}$$

If the slope $\beta > 0$, then the width of the beam solution will be exponentially increasing.

As we can see, the beam solution will lose its accuracy in the simple linear velocity, and each smooth velocity can be approximated by a linear function locally, therefore, the same phenomenon can be expected in other situations.

4.7.2 The Second Motivation for Developing a New Reinitialization Strategy

The second motivation is to resolve the problem caused by reflection beams. The idea in Section 4.5 we have used to derive the reflection formula is theoretically correct, however, it will cause some problems when implementing it numerically, especially in the S-wave reflection case. The difference is that the SP-wave is more likely to be a grazing beam.

To see this, we employ the 2D model for the illustration and let's suppose the ray hitting the line $\{\boldsymbol{x} = (x, y) : x = 0\}$. Then, according to the analysis in Section 4.5, the y-component of the ray direction τ_y does not change. For the PS reflection,

$$
\begin{aligned}
(c^S \tau_x^{PS})^2 &= (c^P |\tau_x^P|)^2 + \left((c^P)^2 - (c^S)^2\right) \tau_y^2 \\
(\tau_x^{PS})^2 &= \frac{(\lambda + 2\mu)|\tau_x^P|^2 + (\lambda + \mu)\tau_y^2}{(c^S)^2},
\end{aligned}
\tag{4.89}
$$

while the SP reflection follows,

$$
(\tau_x^{SP})^2 = \frac{\mu |\tau_x^S|^2 - (\lambda + \mu)\tau_y^2}{(c^P)^2}. \tag{4.90}
$$

We then compute the angle θ^{PS} and θ^{SP} between the ray direction and the reflecting boundary $\{\boldsymbol{x} = (x, y) : x = 0\}$,

$$
\begin{aligned}
\tan(\theta^{PS}) &= \frac{\tau_x^{PS}}{\tau_y^{PS}} \\
&= \sqrt{\frac{(\lambda + 2\mu)|\tau_x^P|^2 + (\lambda + \mu)\tau_y^2}{(c^S)^2 \tau_y^2}} \\
&= \sqrt{\frac{\lambda + \mu}{\mu} + \frac{(\lambda + 2\mu)|\tau_x^P|^2}{\mu \tau_y^2}}.
\end{aligned}
\tag{4.91}
$$

Similarly,

$$
\tan(\theta^{SP}) = \sqrt{\frac{\mu |\tau_x^S|^2}{(\lambda + 2\mu)\tau_y^2} - \frac{\lambda + \mu}{\lambda + 2\mu}}. \tag{4.92}
$$

We claim that for the PS reflection, the angle between the ray direction and the boundary will be increasing after reflection, while for the SP reflection, this value will be decreasing. To see this,

$$\left(\frac{\tan(\theta^{PS})}{\tan(\theta^{P})}\right)^2 = \frac{\frac{\lambda+\mu}{\mu} + \frac{(\lambda+2\mu)|\tau_x^P|^2}{\mu\tau_y^2}}{\left(\frac{\tau_x^P}{\tau_y}\right)^2}$$
$$= \frac{\lambda+2\mu}{\mu} + \frac{\lambda+\mu}{\mu}\left(\frac{\tau_y}{\tau_x^P}\right)^2. \tag{4.93}$$

Apply the same idea to the SP reflection,

$$\left(\frac{\tan(\theta^{SP})}{\tan(\theta^{S})}\right)^2 = \frac{\mu}{\lambda+2\mu} - \frac{\lambda+\mu}{\lambda+2\mu}\left(\frac{\tau_y}{\tau_x^S}\right)^2. \tag{4.94}$$

As we can see from equation (4.93) and equation (4.94), as τ_y is increasing or the incidence angle is decreasing, the ratio for the PS reflection is increasing, which means that the angle after reflection is larger than the incidence angle, while the angle for the SP reflection is decreasing as a quadratic function of τ_y. It means that the angle for the SP wave θ^{SP} will be closer to zero even when the incoming S-wave's incidence angle θ is away from zero. The grazing beam with larger width will interact with the boundary. Therefore, it is needed to be sharpened to guarantee the accuracy.

To remark, there's no problem with the reflection in the acoustic wave as the velocity is the same. We have conducted the experiment to justify the analysis in 3-D space and we show the result in the next section.

4.7.3 Sharpened Wavepackets and Convergence Analysis

By taking advantage of Lemma 4.6.2, we have

$$u(y) = aAe^{i\tilde{B}}\sqrt{\frac{\omega}{2\pi}}e^{-\frac{I}{2}|y-y_0^*|^2}e^{i\tilde{\tau}_y(y-y_0)}e^{i\frac{R}{2}(y-y_0)^2}\int e^{-\omega|y-y_0-ma|^2}dm + O\left(\frac{1}{\omega}\right), \tag{4.95}$$

The term $\tilde{\tau}_y$ is the modified y-direction of the central ray, that is

$$\tilde{\tau}_y = \tau_y - R \cdot E, \tag{4.96}$$

To sharpen Gaussian beams, we have the following lemma,

Lemma 4.7.1.

$$u(y) \approx Ae^{i\tilde{B}}\sum_{k=0}^{q} l_k a\sqrt{\frac{\omega}{2\pi}}e^{i\tilde{\tau}_y(y-y_0)}e^{-\omega_k|y-y_0^*|^2}e^{-\frac{I}{2}|y-y_0^*|^2}e^{i\frac{R}{2}(y-y_0)^2}. \tag{4.97}$$

where $a = \frac{1}{\sqrt{I}}$. Parameters ω_k and q will be given in the proof.

As we can see the extra term $e^{-\omega_k|y-y_0^*|^2}$ reduces the size of beams. And positive ω_k is obtained by choosing parameter ω appropriately.

Proof. To obtain equation (4.97), we first

$$\begin{aligned}\int e^{-\omega|y-y_0-ma|^2}dm &= \int e^{-\omega|y-y_0^*+y_0^*-y_0-ma|^2}dm \\ &= \int e^{-\omega|y-y_0^*|^2}e^{-\omega|y_0^*-y_0-ma|^2}e^{-2\omega(y-y_0^*)(y_0^*-y_0-ma)} \\ &= e^{-\omega|y-y_0^*|^2}\int e^{-\omega|y_0^*-y_0-ma|^2}e^{-2\omega(y-y_0^*)(y_0^*-y_0-ma)}dm.\end{aligned} \tag{4.98}$$

The integral part above is

$$\begin{aligned}&\int e^{-\omega|y_0^*-y_0-ma|^2}e^{-2\omega(y-y_0^*)(y_0^*-y_0-ma)}dm = \\ &\sum_{\tilde{k}}\int_{\tilde{k}-\frac{1}{2}}^{\tilde{k}+\frac{1}{2}} e^{-\omega|y_0^*-y_0-ma|^2}e^{-2\omega(y-y_0^*)(y_0^*-y_0-ma)}dm \\ &= \sum_{k\in\mathbb{Z}}\int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega|(k+\delta)a|^2}e^{2\omega(y-y_0^*)((k+\delta)a)}d\delta.\end{aligned} \tag{4.99}$$

We require k to be integers, which means the value of $\tilde{k}$ satisfies $k = \tilde{k} + \frac{y_0-y_0^*}{a}$. Notice first that we can truncate the above summation to finite terms, $|k| \leq q$, since $e^{-\omega|y-y_0-ma|^2}$ is a L_1 function. By monotone convergence theorem,

$$\lim_{N\to\infty}\int_N^{\infty} e^{-\omega|y-y_0-ma|^2}dm = 0 \tag{4.100}$$

.

Case 1: $k = 0$

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega|(k+\delta)a|^2} e^{2\omega(y-y_0^*)((k+\delta)a)} d\delta = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega|\delta a|^2} e^{2\omega(y-y_0^*)(\delta a)} d\delta$$
$$\approx \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega|\delta a|^2} \left(\sum_n \frac{2^n \omega^n (y-y_0^*)^n (\delta a)^n}{n!} \right) d\delta$$
$$\approx \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega|\delta a|^2} \left(\sum_n \frac{(2\omega)^{2n} (y-y_0^*)^{2n} (\delta a)^{2n}}{(2n)!} \right) d\delta \tag{4.101}$$

The odd power terms in δ vanish in the last step above since $e^{-\omega|\delta a|^2}$ is an even function about δ and all odd power functions are odd functions. The integral of all odd functions in $[-\frac{1}{2}, \frac{1}{2}]$ will be zero.

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega|\delta a|^2} e^{2\omega(y-y_0^*)(\delta a)} d\delta \approx \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega|\delta a|^2} \left(1 + 2\omega^2 (y-y_0^*)^2 (\delta a)^2 \right) d\delta$$
$$\approx \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega|\delta a|^2} e^{2\omega^2 (y-y_0^*)^2 (\delta a)^2} d\delta. \tag{4.102}$$

The leading order error from equation (4.102) is $O\left(\frac{(2\omega)^4 (y-y_0^*)^4 (\delta a)^4}{4!}\right)$.

Case 2: $k \neq 0$

When $k < 0$,

$$\begin{aligned}\int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega|(k+\delta)a|^2} e^{2\omega(y-y_0^*)((k+\delta)a)} d\delta &= \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega|-|k|a+\delta a|^2} e^{2\omega(y-y_0^*)(-|k|a+\delta a)} d\delta \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega|-|k|a-\tilde{\delta}a|^2} e^{2\omega(y-y_0^*)(-|k|a-\tilde{\delta}a)} d\tilde{\delta} \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega||k|a+\delta a|^2} e^{-2\omega(y-y_0^*)(|k|a+\delta a)} d\delta,\end{aligned} \tag{4.103}$$

by setting $\tilde{\delta} = -\delta$.

When $k > 0$

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega|(k+\delta)a|^2} e^{2\omega(y-y_0^*)((k+\delta)a)} d\delta = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega|(|k|+\delta)a|^2} e^{2\omega(y-y_0^*)((|k|+\delta)a)} d\delta \tag{4.104}$$

Given $k > 0$, we add (4.104) for $k > 0$ and (4.103) for $-k < 0$, so that we have

$$\begin{aligned}&\int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega|ka+\delta a|^2} \left(e^{2a\omega(y-y_0^*)(k+\delta)} + e^{-2a\omega(y-y_0^*)(k+\delta)}\right) d\delta \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} 2e^{-\omega|ka+\delta a|^2} \left(1 + 2\omega^2(y-y_0^*)^2(k+\delta)^2a^2\right) d\delta\end{aligned} \tag{4.105}$$

Here we use the Taylor expansion of the exponential function. Furthermore,

$$\begin{aligned}&\int_{-\frac{1}{2}}^{\frac{1}{2}} 2e^{-\omega|ka+\delta a|^2} \left(1 + 2\omega^2(y-y_0^*)^2(k+\delta)^2a^2\right) d\delta \\ &\approx \int_{-\frac{1}{2}}^{\frac{1}{2}} 2e^{-\omega(k+\delta)^2a^2} e^{2\omega^2(y-y_0^*)^2((k+\delta)a)^2} d\delta.\end{aligned} \tag{4.106}$$

The error term is $O\left(\frac{(2\omega)^4(y-y_0^*)^4(\delta a+ka)^4}{4!}\right)$. The error of approximations (4.106) and (4.102) can be summarized as an universe form, i.e, $O\left(\frac{(2\omega)^4(y-y_0^*)^4(\delta a+ka)^4}{4!}\right)$.

Now integral (4.99) becomes

$$\begin{aligned}
&\int e^{-\omega|y_0^*-y_0-ma|^2}e^{-2\omega(y-y_0^*)(y_0^*-y_0-ma)}dm = \\
&\int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-\omega||k|a+\delta a|^2}e^{-2\omega(y-y_0^*)(|k|a+\delta a)}d\delta \\
&+\sum_{k>0}\int_{-\frac{1}{2}}^{\frac{1}{2}} 2e^{-\omega(k+\delta)^2a^2}e^{2\omega^2(y-y_0^*)^2((k+\delta)a)^2}d\delta.
\end{aligned} \tag{4.107}$$

To get the value l_k and ω_k, we start with $k>0$,

$$\begin{aligned}
&e^{-\omega|y-y_0^*|^2}\int_{-\frac{1}{2}}^{\frac{1}{2}} 2e^{-\omega(k+\delta)^2a^2}e^{2\omega^2(y-y_0^*)^2((k+\delta)a)^2}d\delta \approx \\
&2e^{-\omega|y-y_0^*|^2}e^{2\omega^2(y-y_0^*)^2(ka)^2}e^{-\omega|ka|^2} \\
&= l_k e^{-\omega|y-y_0^*|^2}e^{2\omega^2(y-y_0^*)^2(ka)^2}
\end{aligned} \tag{4.108}$$

First, the first step above is obtained by choosing $\delta = 0$. Second, the value of l_k is defined by

$$l_k = 2e^{-\omega|ka|^2}, \tag{4.109}$$

Similarly, when k = 0,

$$l_k = 1. \tag{4.110}$$

The next step is to compute the value of ω_k and define the parameter ω. We start with making the exponent in equation (4.108) negative,

$$
\begin{aligned}
-\omega + 2\omega^2 k^2 a^2 &\leq 0 \\
\omega &\leq \frac{1}{2a^2k^2}, \quad k = 1, 2, \cdots, q
\end{aligned}
\tag{4.111}
$$

and a is previously defined as $\frac{1}{\sqrt{I}}$ in Lemma 4.6.2. If we choose $\omega = \frac{I}{3q^2}$, then ω_k in equation (4.97) is

$$
\omega_k = \left(\frac{3q^2 - 2k^2}{9q^4} \right) I. \tag{4.112}
$$

After defining the values l_k and ω_k, we will characterize the size of the error term,

$$
\begin{aligned}
\max_{k \leq q} \frac{(2\omega)^4 (y - y_0^*)^4 (\delta a + ka)^4}{4!} &= \frac{(2\omega)^4 (y - y_0^*)^4 (\delta a + qa)^4}{4!} \\
&\leq \frac{2}{3} \left(\frac{I}{3q^2} \right)^4 \left(\frac{1}{\sqrt{I}} \right)^4 (qa)^4 \\
&\leq \frac{2}{243} \frac{1}{q^4}
\end{aligned}
\tag{4.113}
$$

Although this is only the leading order in the series we truncated, the rest of them will be easily controlled from the following two concerns. The first concern is the rest of the series will have high power about q, i.e. $\frac{1}{q^{2n}}$, for $n = 3, 4, \cdots$. The second concern is that the coefficient $\frac{2^{2n}}{(2n)!}$ will decay exponentially. $\square$

4.8 Numerical Examples

In this section, we present all numerical experiments conducted to justify the proposed method through some complex velocity and general initial conditions. All the reference solution in this section are provided by the FDTD method with staggered grid. Reference solution's correctness will be examined in Appendix B.

4.8.1 Beam Reinitialization

4.8.1.1 SP Reflection V.S. PS Reflection

In the subsection, we will provide the numerical results of the analysis in Section 4.7.2 and Section 4.7.3. We firstly conduct the following experiment to illustrate the difference between the SP reflection and the PS reflection. The experiment's setup for the S-wave is,

$$\boldsymbol{f}^S = \begin{pmatrix} -\sin(\alpha)\sin(36\pi x + 36\pi\sin(\alpha)(y-0.25))e^{-36\pi^2\left(x^2+(y-0.25)^2+(z-0.25)^2\right)} \\ \sin(36\pi x + 36\pi\sin(\alpha)(y-0.25))e^{-36\pi^2\left(x^2+(y-0.25)^2+(z-0.25)^2\right)} \\ 0 \end{pmatrix}, \tag{4.114}$$

and

$$\boldsymbol{g}^S = \boldsymbol{0}. \tag{4.115}$$

The experiment's setup for P-wave is,

$$\boldsymbol{f}^P = \begin{pmatrix} \sin(36\pi x + 36\pi \sin(\alpha)(y-0.25))e^{-36\pi^2\left(x^2+(y-0.25)^2+(z-0.25)^2\right)} \\ \sin(\alpha)\sin(36\pi x + 36\pi \sin(\alpha)(y-0.25))e^{-36\pi^2\left(x^2+(y-0.25)^2+(z-0.25)^2\right)} \\ 0 \end{pmatrix},$$

and

$$\boldsymbol{g}^P = \mathbf{0} \tag{4.116}$$

The initial width of the beams are all set to be $36\pi^2$ and $\lambda = \mu = 1$. The experiment's result is displayed in Figure 4.4. The red star line in Figure 4.4 is the ratio $\frac{\sin(\theta^{PS})}{\sin(\alpha)}$,

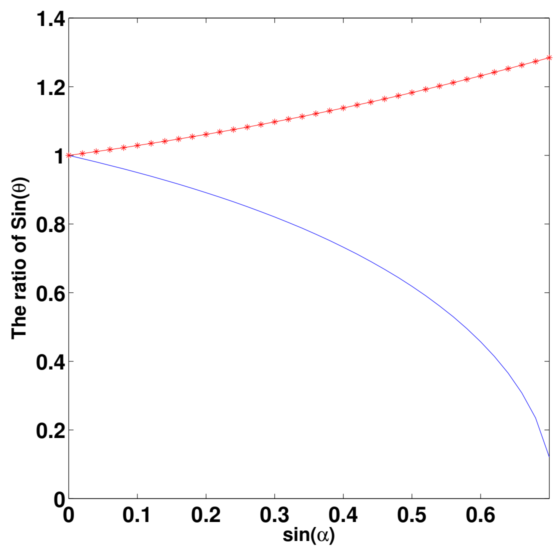


Figure 4.4: PS Reflection V.S. SP Reflection: Different Ratio Behaviors.
PS Reflection('o'), SP Reflection ('-')

and the blue line is the ratio $\frac{\sin(\theta^{SP})}{\sin(\alpha)}$. Same as the analysis in Section 4.7.2, the angle for the SP reflection will decrease to zero compared with the original hitting angle. Therefore, the SP reflection wave should be modified to be more focused.

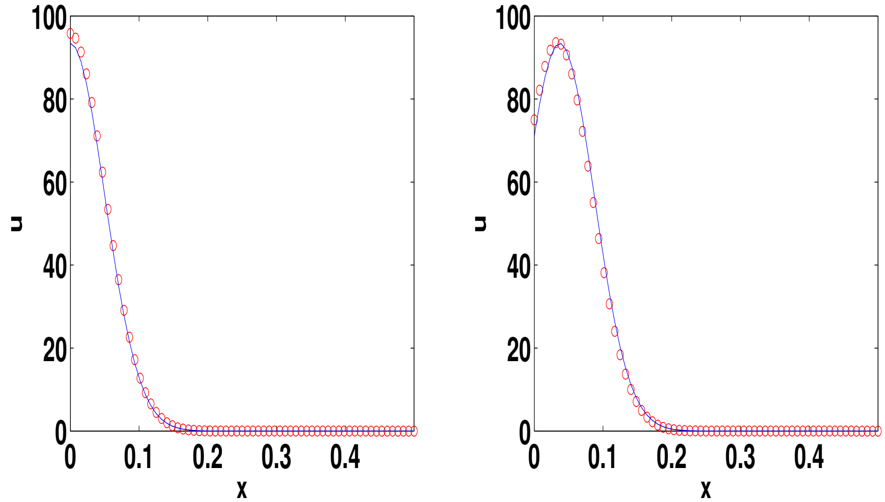


Figure 4.5: Sharpened Beams V.S. the Original Beam on Fixed (y, z).
Left: y=0.1, z=0.1 **Right:** y=0.08, z=0.13.
Beam Solution after Reinitialization('o'), Original Beam Solution ('-').

4.8.1.2 Sharpened Beams V.S. Original Beams

The second experiment in this subsection is conducted to show the fact that the precision will not be impacted after adding the new reinitialization (4.97) process. We compare the wavefield after the reinitialization with the original one. Suppose the S-wave hits at the boundary $x = 0$ with the central point $(0, 0.1, 0.1)$ and the Hessian equals to

$$\begin{pmatrix} 36\pi + 36\pi^2 i & 12\pi^2 i & 0 \\ 12\pi^2 i & 7\pi^2 + 36\pi^2 i & 0 \\ 0 & 0 & 4\pi^2 + 36\pi^2 i \end{pmatrix}.$$

Its amplitude norm is set as 100. We choose two different sets of the y-value and z-value to show our reinitialization method's correctness, The wavefield in Figure 4.5 is plotted along the x-axis. As we can see, the reinitialization algorithm won't affect the accuracy of the result, while the width of each new beam has been decreased.

4.8.2 Periodic Boundary Condition

In this subsection, we would like to display some numerical results to the periodic boundary problem.

4.8.2.1 Example 1: The Single Wavepacket in the Constant velocity

We first test our algorithm via the constant elastic moduli $\lambda = \mu = 1$, and the nonzero initial velocity $\mathbf{g}$.

$$\boldsymbol{f} = \begin{pmatrix} \sin(36\pi x + 9\pi y)e^{-36\pi^2\left((x-0.25)^2+(z-0.25)^2+(z-0.25)^2\right)} \\ \sin(36\pi x + 9\pi y)e^{-36\pi^2\left((x-0.25)^2+(z-0.25)^2+(z-0.25)^2\right)} \\ 0 \end{pmatrix}, \tag{4.117}$$

and the initial velocity $\boldsymbol{g}$ is

$$\boldsymbol{g} = 0.5\boldsymbol{f}. \tag{4.118}$$

We exhibit our result at the plane $z = 0.25$. The comparison between two results shows that the propagation dynamics, including the eikonal equation and the transport equation are correct. The Multi-scale Gaussian Wavepacket Transform is also justified, although this is a single wavepacket initial value. We will test the transform further with some more general initial condition.

4.8.2.2 Example 2: General Initial Condition in the Constant velocity

The next thing we would like to try is some general initial condition other than the single Gaussian wavepacket to verify our initial decomposition. The initial value is

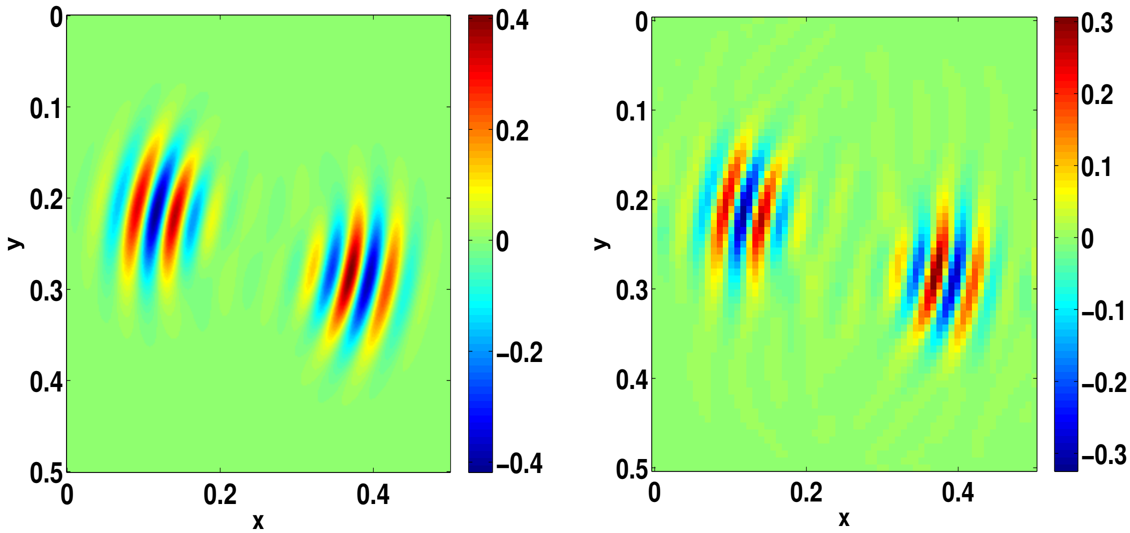


Figure 4.6: **Example 1:**
Single Wavepacket Propagation with the Periodic Boundary Condition.
Left: FDTD Solution **Right:** Gaussian Beam Solution.

defined as

$$\boldsymbol{f} = \begin{pmatrix} \sin(72\pi x^2)e^{-36\pi^2\left((x-0.25)^2+(z-0.25)^2+(z-0.25)^2\right)} \\ \sin(72\pi x^2)e^{-36\pi^2\left((x-0.25)^2+(z-0.25)^2+(z-0.25)^2\right)} \\ 0 \end{pmatrix} \tag{4.119}$$

and

$$\boldsymbol{g} = 0; \tag{4.120}$$

Let's compare the result along the x-axis by setting $z = 0.25$ and $y = \frac{6}{32}$. Now let's compare 2-D wavefields at $z = 0.25$.

For other more complex velocity, we will exhibit those results in the Dirichlet boundary condition, since the periodic boundary problem essentially tests the correctness of propagation dynamic and the decomposition process of the initial condition as proposed

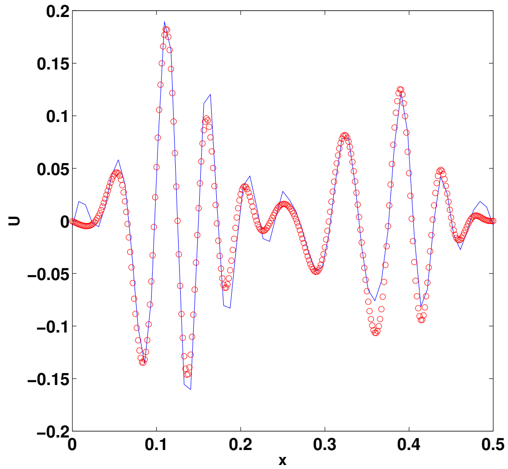


Figure 4.7: **Example 2:**
General Initial Value Propagation with Periodic Boundary Condition along x-axis
FDTD Solution('o'), Gaussian Beam Solution ('-')

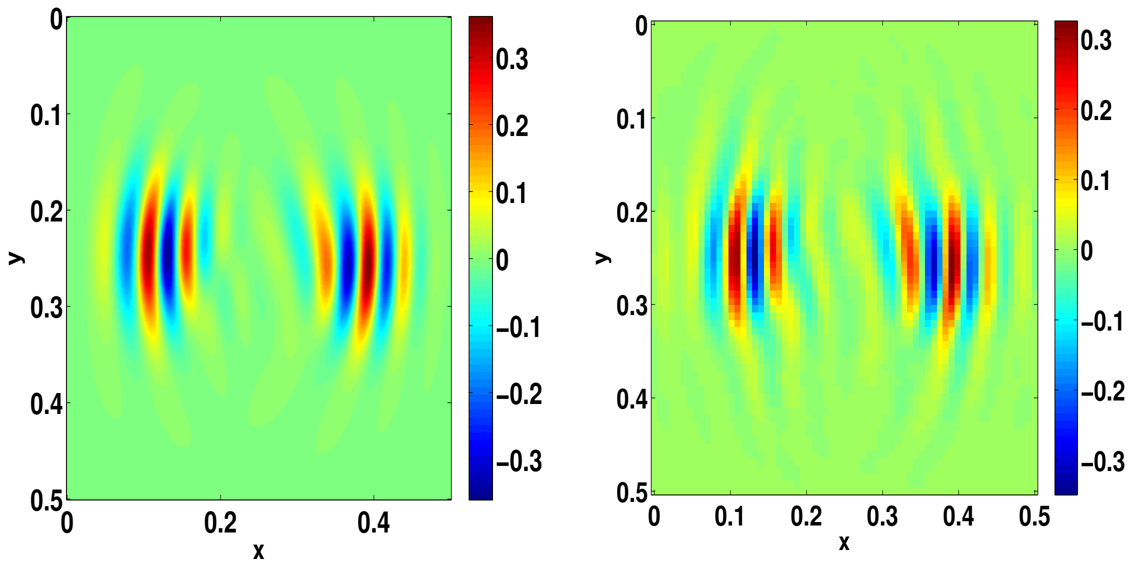


Figure 4.8: **Example 2:**
General Initial Value Propagation with Periodic Boundary Condition.
Left: FDTD Solution **Right:** Gaussian Beam Solution.

in Section 4.4. This can be examined in the Dirichlet boundary problem as well.

4.8.3 Reflection: P-wave

We have so far used the periodic boundary condition to justify our decomposition process and the dynamic system, now we would like to test our reflection scheme on the pure P-wave initial condition. In this subsection, we fix our initial condition as the following and test our algorithms over the different velocity.

$$\boldsymbol{f} = \begin{pmatrix} \sin(32\pi x + 8\pi y)e^{-36\pi^2\left((x-0.15)^2+(y-0.25)^2+(z-0.25)^2\right)} \\ \frac{1}{4}\sin(32\pi x + 8\pi y)e^{-36\pi^2\left((x-0.15)^2+(y-0.25)^2+(z-0.25)^2\right)} \\ 0 \end{pmatrix} \tag{4.121}$$

and the initial velocity $\boldsymbol{g}$ is

$$\boldsymbol{g} = \begin{pmatrix} 2\sin(32\pi x + 8\pi y)e^{-36\pi^2\left((x-0.15)^2+(y-0.25)^2+(z-0.25)^2\right)} \\ \frac{1}{2}\sin(32\pi x + 8\pi y)e^{-36\pi^2\left((x-0.15)^2+(y-0.25)^2+(z-0.25)^2\right)} \\ 0 \end{pmatrix} \tag{4.122}$$

4.8.3.1 Example 3: P-wave Reflection in the Constant velocity

Our first setting is still under the constant elastic moduli, $\lambda = 1$ and $\mu = 2$. We compare the wavefield at $z = 0.25$ and $T = 0.14$, after the primary reflection happens. The P-wave reflection dynamics can be justified after this numerical experiment.

Here the mesh size of the FDTD method employed is $\frac{1}{640}$ and this scale will be used for all the FDTD results in the rest of the paper.

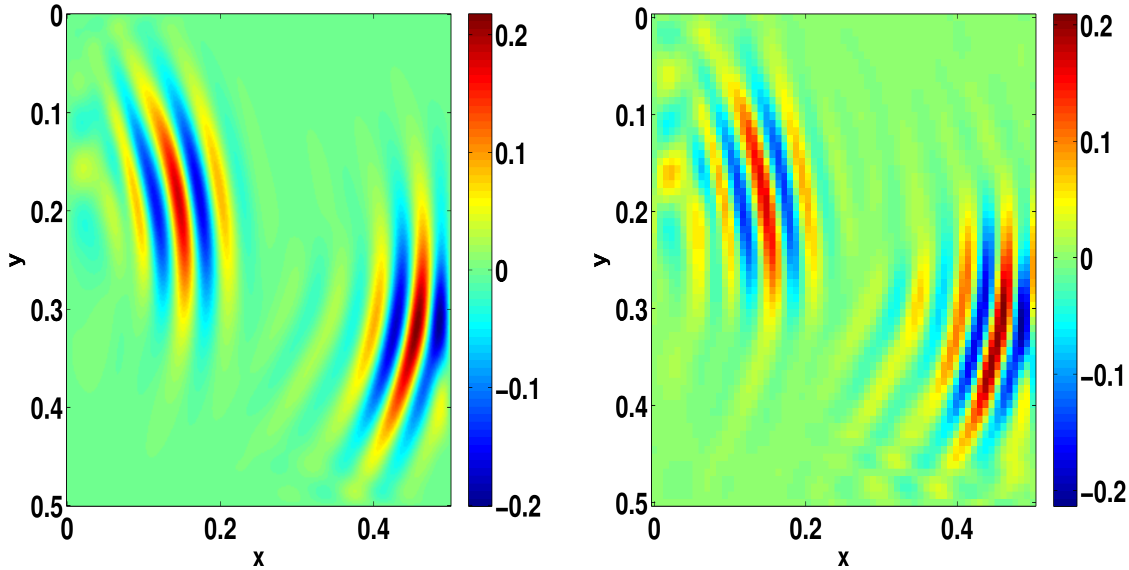


Figure 4.9: **Example 3:**
P-wave Reflection in Constant velocity with Dirichlet Boundary Condition.
Left: FDTD Solution **Right:** Gaussian Beam Solution.

4.8.3.2 Example 4: P-wave Reflection in the Linear velocity

The second velocity in the P-wave reflection section we use is $\mu = 2$ and $\lambda = 1 + 0.2x$. We first compare the result at the fixed $(y, z) = (0.125, 0.25)$ along x-axis. Then we set z fixed as 0.25, The P-wave reflection dynamics in some general velocity has been justified further. As we can see in Figure 4.10, the Gaussian beam solution performs quite well in the major region, and the tolerable error shows up near the boundary.

4.8.3.3 Example 5: P-wave Reflection in the Sinusoidal velocity

Now let's try the sinusoidal elastic moduli $\lambda = 1 + \sin(4\pi x)$, in which there will be some caustics points. Again, we compare the result along the x-axis first by fixing $y = z = 0.25$.

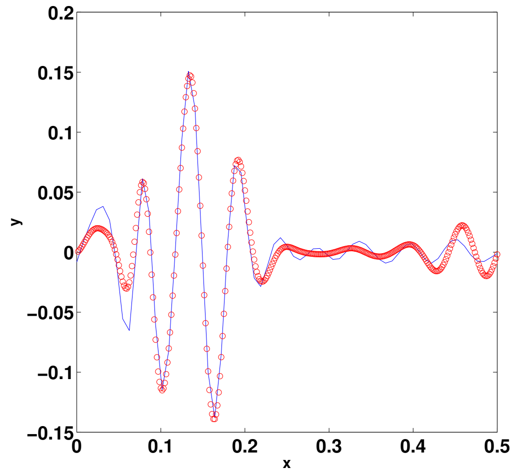


Figure 4.10: **Example 4:**
P-wave Reflection in the Linear velocity with Dirichlet Boundary Condition.
FDTD Solution ('o'), Beam Solution('-')

Then we fix $z = 0.25$, To remark, the reinitialization scheme is not involved in all the P-wave reflection results shown above. The small error implies that the P-wave reflection does not require the reinitialization, while it is necessary in the S-wave reflection.

4.8.4 Reflection S-wave

In this subsection, we will test the reinitialization process, but at first we will justify the necessity of adding reinitialization process.

4.8.4.1 Example 6: S-wave Reflection with Orthogonal Hitting Angle

Firstly, we see the S-wave reflection with orthogonal hitting angle, that is $\sin(\alpha) = 0$ in Figure 4.4. It is displayed by Figure 4.14 showing that the original method without the reinitialization is good enough.

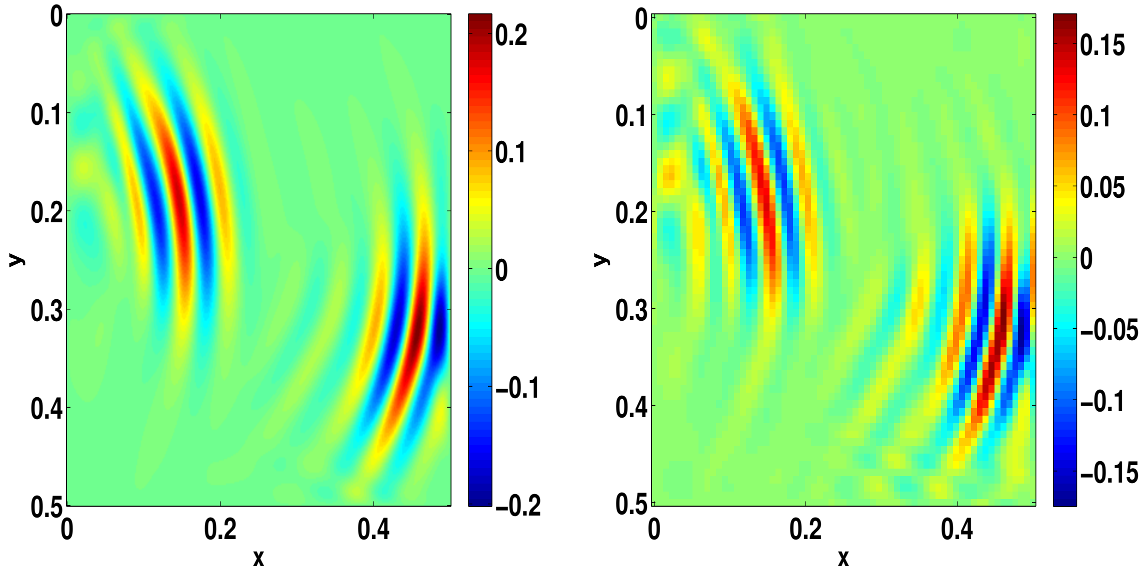


Figure 4.11: **Example 4:**
P-wave Reflection in Linear velocity with Dirichlet Boundary Condition.
Left: FDTD Solution **Right:** Gaussian Beam Solution.

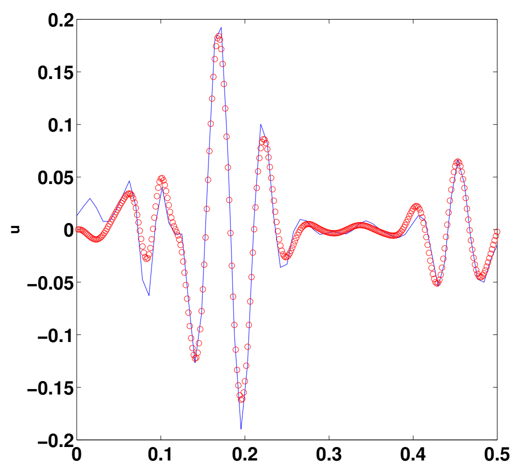


Figure 4.12: **Example 5:**
P-wave Reflection in Sinusoidal velocity along x-axis.
FDTD Solution('o'), Beam Solution ('-')

The setup of our experiment is

$$\boldsymbol{f} = \begin{pmatrix} \sin(36\pi y)e^{-36\pi^2((x-0.25)^2+(y-0.15)^2+(z-0.25)^2)} \\ 0 \\ 0 \end{pmatrix}. \tag{4.123}$$

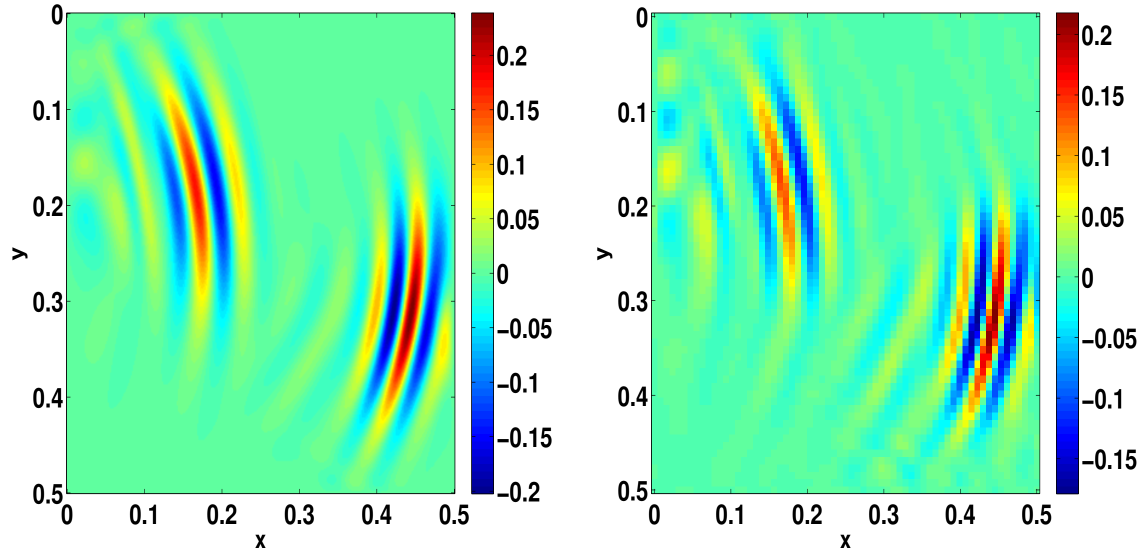


Figure 4.13: **Example 5:**P-wave Reflection in Sinusoidal velocity with Dirichlet Boundary Condition. **Left:** FDTD Solution, **Right:** Gaussian beam Solution

and there is no initial velocity, i.e. $\boldsymbol{g} = 0$. Two velocity parameters λ and μ here are both constants,

$$\lambda = 1; \mu = 2;$$

Figure 4.14 compares the wave field generated by our method to the one from FDTD method at $z = 0.25, x = 0.25$ and $T = 0.2$. Now we fix $z = 0.25$ and $y = \frac{5}{32}$, As we can see from Figure 4.4, when $\sin(\alpha) = 0$, the regular reflection method is expected to be well enough and the experiment result above justifies our conclusion.

4.8.4.2 Example 7: S-wave Reflection with Non-Orthogonal Hitting Angle

Now if we change the ray direction to increase the width of the SP-wave, we will see that the regular reflection dynamics fails in this case. To make that happen, we specify

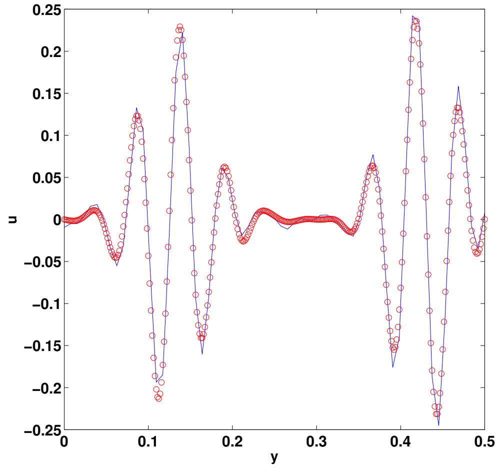


Figure 4.14: **Example 6:**
S-wave Reflection with Orthogonal hitting Angle along y-axis
FDTD Solution ('o'), GB Solution ('-').

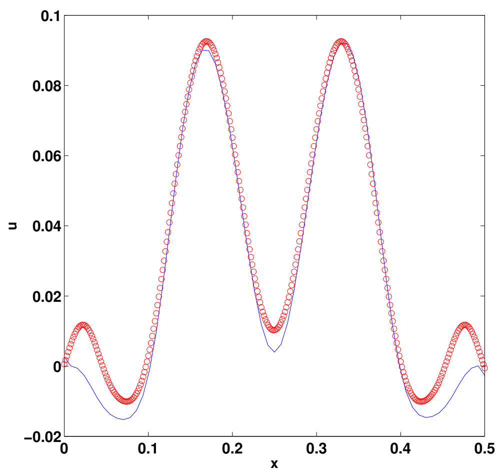


Figure 4.15: **Example 6:** S-wave Reflection with Orthogonal hitting Angle along x-axis. FDTD Solution ('o'), GB Solution ('-').

our initial value as,

$$\boldsymbol{f} = \begin{pmatrix} 2\sin(36\pi y + 18\pi x)e^{-36\pi^2\left((x-0.25)^2+(y-0.15)^2+(z-0.25)^2\right)} \\ -\sin(36\pi y + 18\pi x)e^{-36\pi^2\left((x-0.25)^2+(y-0.15)^2+(z-0.25)^2\right)} \\ 0 \end{pmatrix}. \tag{4.124}$$

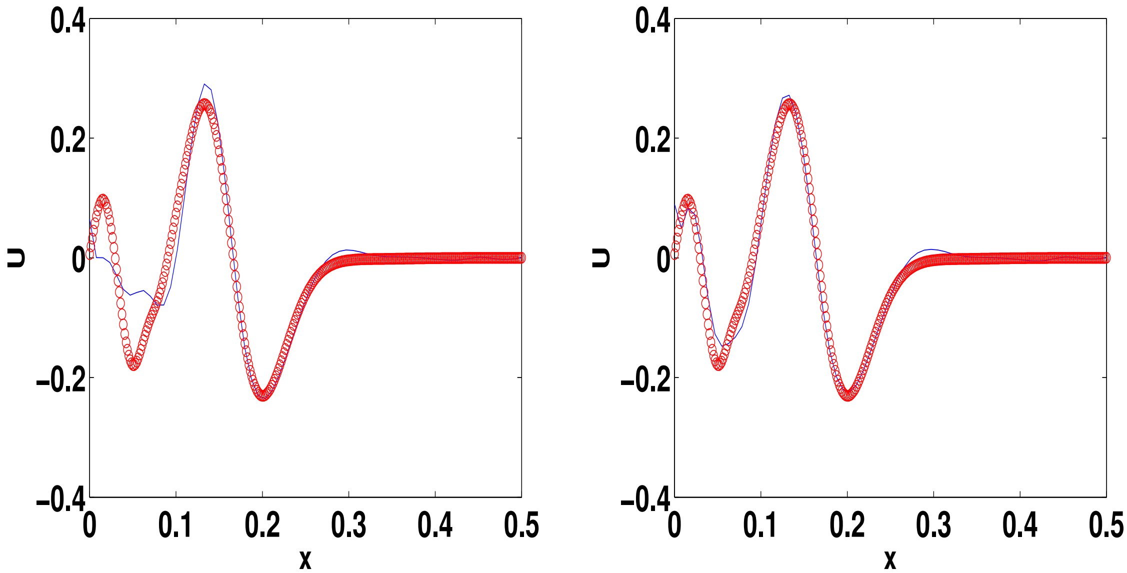


Figure 4.16: **Example 7:** S-wave Reflection with Non-orthogonal Hitting Angle. **Left:** Gaussian Beam Solution without reinitialization, **Right:** Gaussian Beam Solution with reinitialization. FDTD Solution ('o'), GB Solution ('-')

We first see that the result generated without the extra reinitialization. The above left Figure 4.16 is plotted along the x-axis with $z = 0.25, y = 0.125$. As we mentioned before, although the main pattern is captured with good accuracy, the tail region of the Gaussian beam wavefield is not clean enough due to the fact that the SP-wave is involved.

After adding the reinitialization, in the above right Figure 4.16, the beam solution with the reinitialization shows the better result in the tail region without hurting the accuracy of other parts.

The experiment shown in Figure 4.16 illustrates that the analysis in Section 4.7 is correct. The SP-wave reflection will lose the accuracy to some degree such that adding the reinitialization scheme is necessary.

4.8.4.3 Example 8: S-wave Reflection: Linear velocity

Let's see the comparison under some more complicated elastic moduli $\mu = 2 + 0.2y$ and $\lambda = 1$. We compare the wavefield without the reinitialization to the FDTD result first at $z = 0.25$.

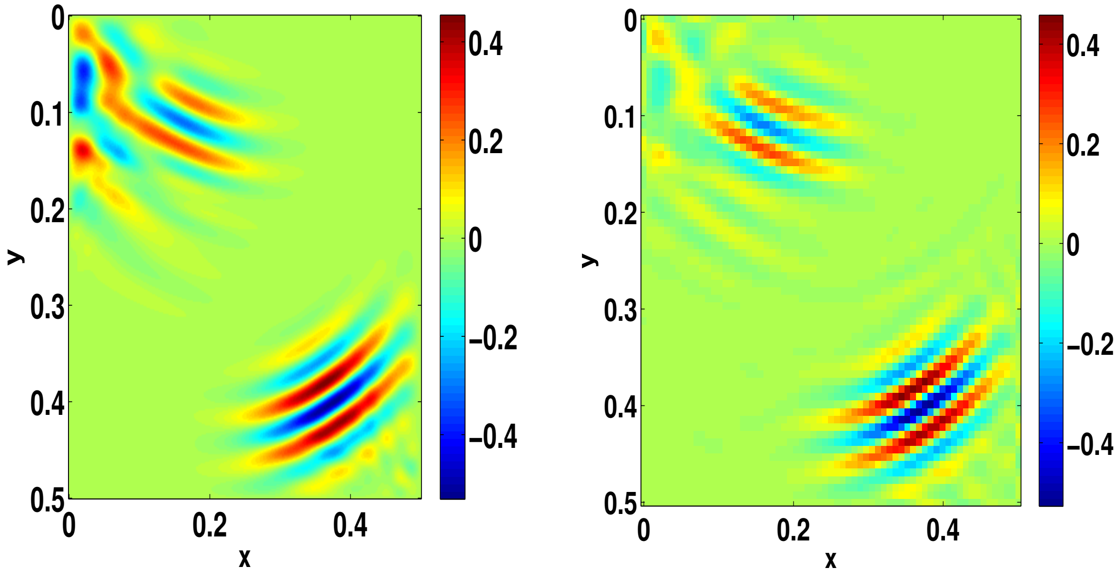


Figure 4.17: **Example 8:** S-wave Reflection in Linear velocity without Reinitialization. **Left:** FDTD Solution **Right:** Gaussian Beam Solution without Reinitialization.

As we can see in the upper left corner of Figure 4.17, there is some significant perturbations in the FDTD result, while the beam method without the reinitialization, which is shown in Figure 4.17, is not able to cover that part.

Now we will see the result after using the reinitialization process. As we can see from the right Figure 4.20, the missing part in the upper left region is covered by the reinitialization algorithm.

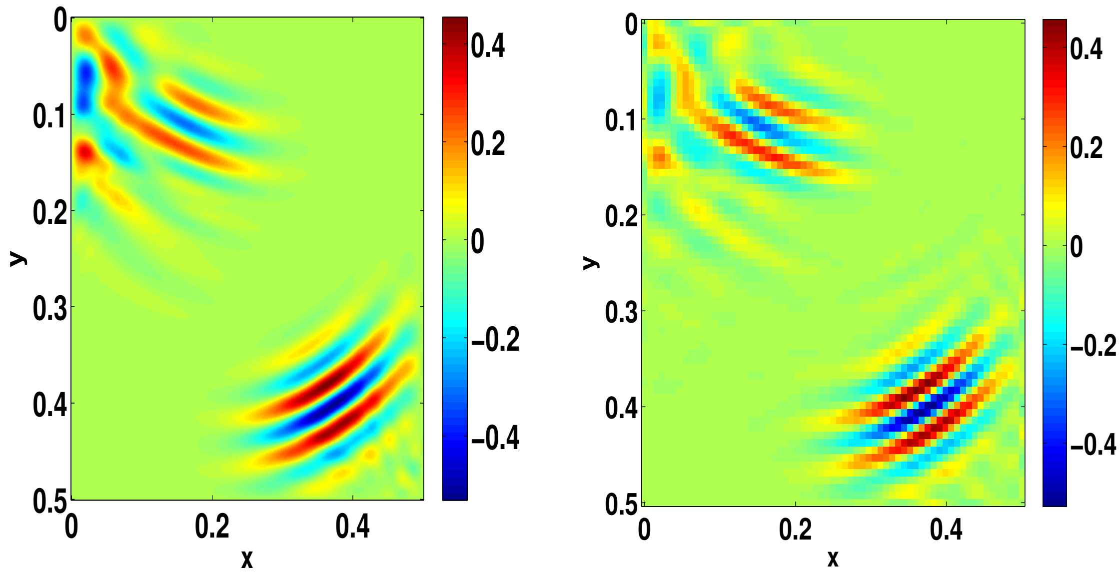


Figure 4.18: **Example 8:** S-wave Reflection in Linear velocity with Reinitialization. **Left:** FDTD Solution **Right:** Gaussian Beam Solution with Reinitialization.

4.8.4.4 Example 9: S-wave Reflection: Sinusoidal velocity

In this example, we set the elastic moduli $\mu = 2 + 0.2\sin(x)$ to make the Hessian of the velocity nonzero. And all the other components remain the same as the last one, including the initial value, λ and the terminal time T.

Compare two wavefields on the plane $z = 0.25$,

In this more complex velocity, the advantage of the Gaussian beam ansatz has shown up as the caustics problem is resolved automatically. There will be caustics in this sinusoidal velocity as the eikonal equation will be multivalued in some region. As Figure 4.20 suggests, the beam solution will perform well even when the caustics shows up.

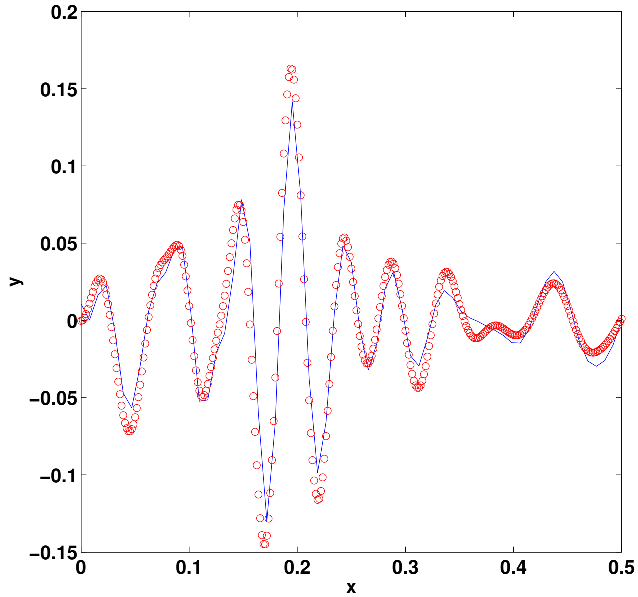


Figure 4.19: **Example 9:** S-wave Reflection in Sinusoidal velocity. GB Solution('o'), FDTD Solution('-')

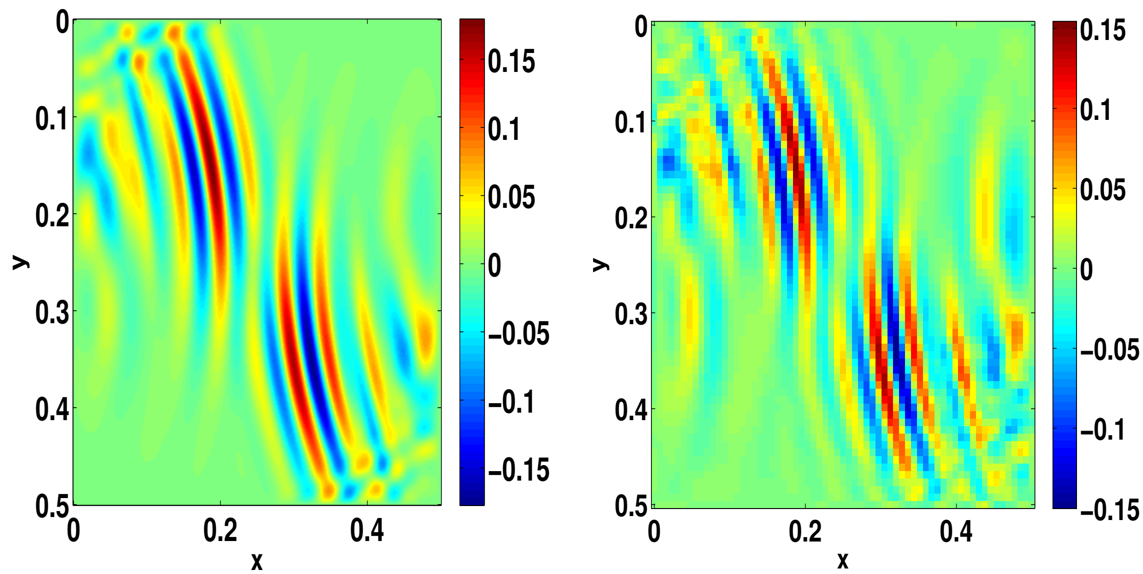


Figure 4.20: **Example 9:** S-wave Reflection in Sinusoidal velocity with Reinitialization. **Left:** FDTD Solution **Right:** Gaussian Beam Solution with Reinitialization.

4.8.5 General Initial Condition

Now after showing the effect of the reinitialization, especially after comparing it to the result without the reinitialization, it is confident to say that the proposed reinitialization algorithm is correct and necessary in this problem. We will end our numerical tests with some more general initial conditions.

4.8.5.1 Example 10: The General Initial Condition in The Sinusoidal velocity

We test Multiscale Gaussian Wavepacket method via the general initial condition and sinusoidal elastic moduli $\lambda = 1 + \sin(4\pi x), \mu = 2$. The initial condition is

$$\mathbf{f} = \begin{pmatrix} \sin(32\pi x + 32\pi y^2) e^{-36\pi^2\left((x-0.15)^2+(y-0.25)^2+(z-0.25)^2\right)} \\ \sin(32\pi x + 10\pi y) e^{-36\pi^2\left((x-0.15)^2+(y-0.25)^2+(z-0.25)^2\right)} \\ \sin(16\pi z) e^{-36\pi^2\left((x-0.15)^2+(y-0.25)^2+(z-0.25)^2\right)} \end{pmatrix}. \tag{4.125}$$

Here we let our wavefield propagate until $T = 0.2$.

We first display the first component in the resulting wavefield. Let's see the second component of the resulting wavefield, and see the comparison between two methods along the x-axis by fixing $y = z = 0.25$. With more general initial value, we include P and S-wave at the same time, meanwhile, different types of the reflection happen simultaneously, so as the different reflection modes.

Remark 4.8.1. *Compared with the FDTD algorithm with parallel computing scheme, our asymptotic algorithm has the larger time complexity. However, as the wavenumber of the initial value is increasing, the FDTD scheme requires finer grid size, leading to the requirement of some larger storage. This is impossible in the current GPU units, while our method's storage complexity is independent of the wavenumber as well as the time complexity.*

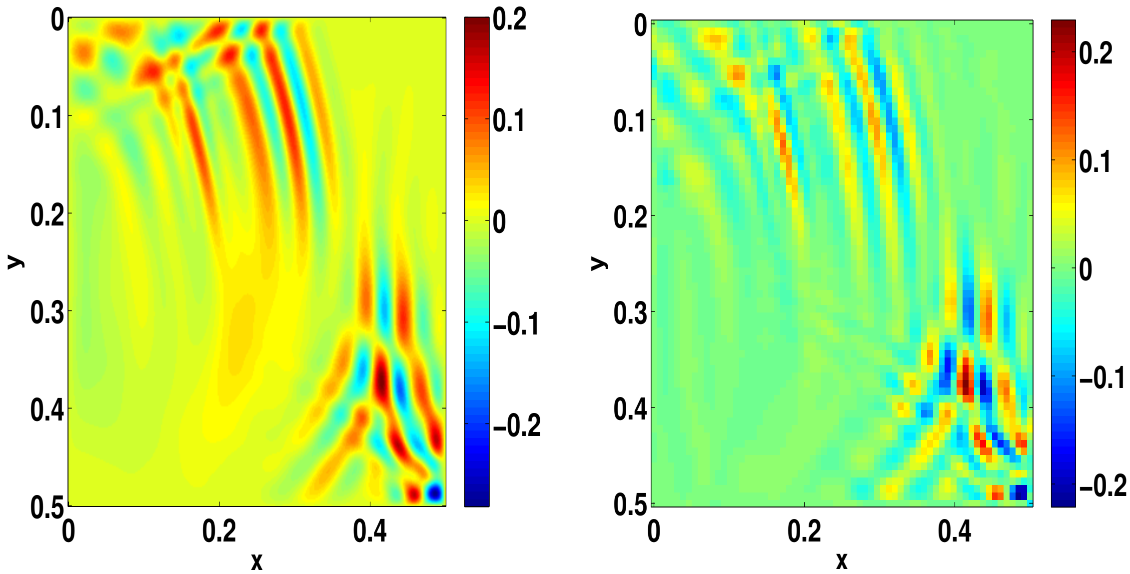


Figure 4.21: **Example 10:** General Initial Condition Propagation with Reinitialization (First Component). **Left:** FDTD Solution **Right:** Gaussian Beam Solution with Reinitialization.

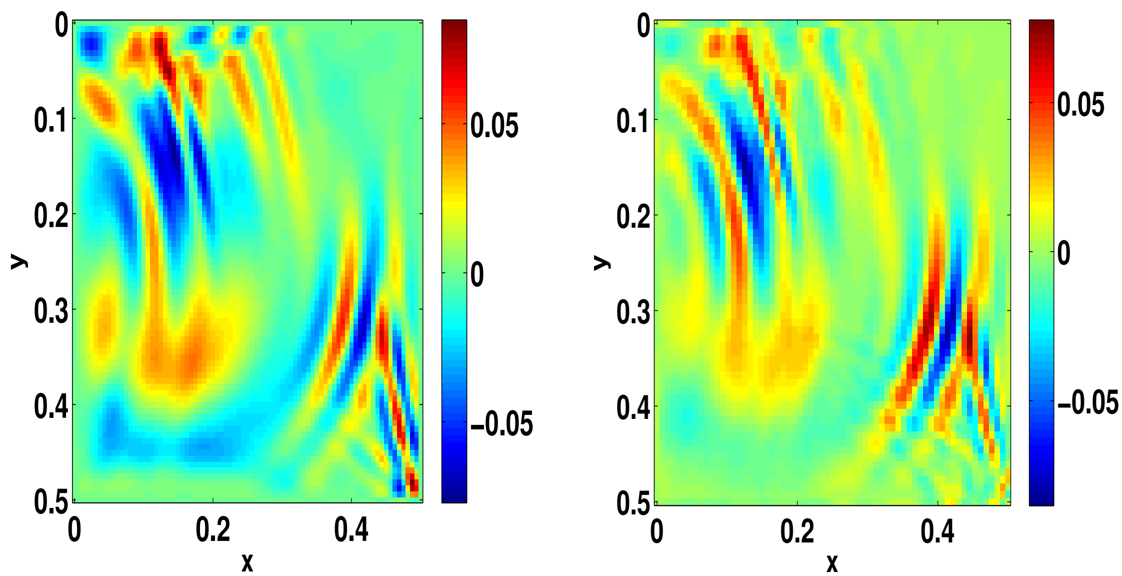


Figure 4.22: **Example 10:**
General Initial Condition Propagation with Reinitialization (Second Component).
Left: FDTD Solution **Right:** Gaussian Beam Solution with Reinitialization.

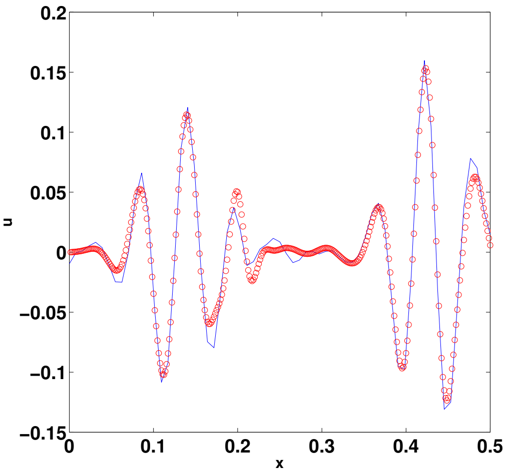


Figure 4.23: **Example 10:** General Initial Condition Propagation with Reinitialization (Second Component). FDTD Solution ('o'), GB Solution('-').

4.8.5.2 Convergence Rate Analysis

In the end, we propose the convergence rate analysis. The initial condition is set as

$$\boldsymbol{f} = \begin{pmatrix} \sin(\eta(16\pi x + 8\pi y))e^{-42\pi^2\left((x-0.25)^2+(y-0.15)^2+(z-0.25)^2\right)} \\ \sin(\eta(16\pi x + 8\pi y))e^{-42\pi^2\left((x-0.25)^2+(y-0.15)^2+(z-0.25)^2\right)} \\ 0 \end{pmatrix} \tag{4.126}$$

The velocity here are all constants,

$$\lambda = 1; \mu = 2;$$

The amplifying factor η is a geometric series, $1, 1.5, 2.25, \cdots, 1.5^5$.

The blue star line is the logarithm of the L_2 norm of the error at different η, while the red line is the linear function with the slope as $\frac{1}{2}\log(1.5)$. It is well known that the convergence rate of the Gaussian beam is $\frac{1}{\sqrt{\omega}}$, and as proved in the paper [5], the

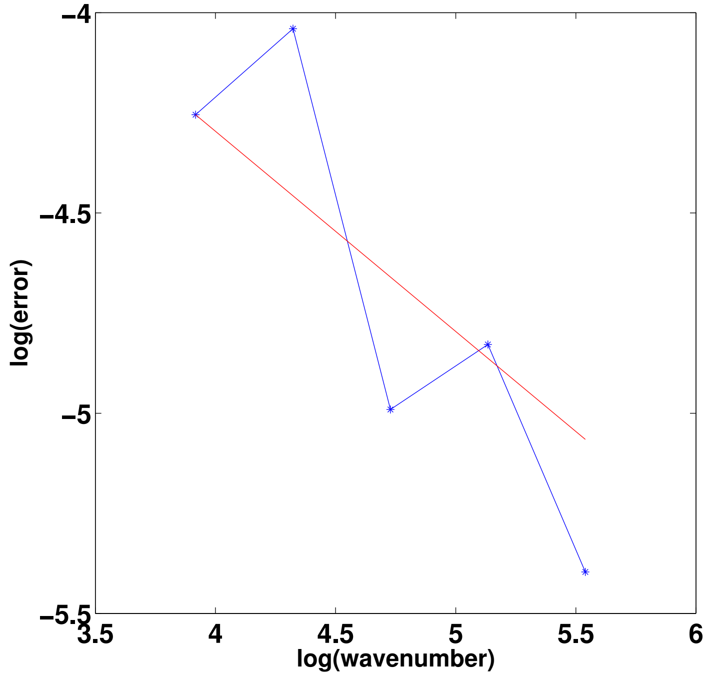


Figure 4.24: **Log-log plot: Convergence Rate of the New Gaussian Beam Method** GB Method Error Curve('-*'), Line with the slope $= \frac{1}{2}\log(1.5)$('-')

Multiscale Gaussian wavepacket transform also follows $O(\frac{1}{\sqrt{\omega}})$. This pattern can be seen in Figure 4.24.

Chapter 5

Conclusion

We propose two methods based on multi-scale Gaussian beam method in this thesis. The first one is solving the elastic wave propagation in the bounded domain and the second one is for the prestack inversion process, which is an inverse problem in geophysical applications. These two methods are both based on the multiscale Gaussian beam method described in Chapter 2. Therefore, both methods are capable of resolving caustics problem automatically and they both take advantage of the parabolic scaling principle for efficiency.

In the first part (Chapter 3), we present a new prestack inversion process, which connects the boundary data to the wavefront set of the perturbation. We first modify the multiscale Gaussian wavepacket transform [48] appropriately to suit to the imaging operator. Secondly, the multi-valued traveltime information is preserved due to the Gaussian beam function. This improves the quality of resulting migration image. Another big advantage of our multiscale Gaussian beam inversion method is its robustness to the polluted data. Since we recover the reflector by its wavefront set, the noise in the boundary data, which is far from the target frequency, won't affect the imaging result. Lastly, our imaging condition is performed in the time domain to avoid the extra Fourier transform on the data set. This feature makes our algorithm more applicable considering the large size of the trace dataset in the real world.

In the second part (Chapter 4), we present a novel Multiscale Gaussian beam method to solve the elastic wave equation in the bounded domain. Firstly, a new vector-valued wavepacket transform is developed to adapt to the highly oscillated vector-valued initial condition following the parabolic scaling principle. Secondly, a novel reinitialization strategy is added in the process to improve the efficiency and accuracy. There are several advantages about this new reinitialization method. The first one is that the new reinitialization is applied to the single wavepacket instead of the whole wavefield. This will improve the efficiency greatly. The second one is that the center of each new wavepacket after the reinitialization is same to the center before, which guarantees all computation happening inside the domain without extra assumption outside the domain. Although the typical FDTD (**F**inite **D**ifference **T**ime **D**omain) method is faster by implementing in parallel, the requirement of large storage will still make FDTD method unfeasible in the high frequency regime.

APPENDICES

Appendix A

Proof in Inverse Process

Hessian Matrix in Corollary 3.3.1

In Corollary 3.3.1, the quadratic term $||r - \hat{r}_0||^2_{\hat{\mathcal{M}}(\hat{t}_c)}$ in the phase function is not necessary to be a Gaussian profile, since $\hat{\mathcal{M}}$ is only the Schur complement of a semi positive definite matrix.

Lemma A.0.1. *For any boundary points $r = \{x = (x_1, \cdots, x_d) : x_d = 0\}$, we have*

$$r^T \left(Im(\hat{M})(\hat{t}_c) - \frac{Im(\hat{\tau}_{tx})(\hat{t}_c) Im(\hat{\tau}_{tx})^T(\hat{t}_c)}{Im(\hat{\tau}_{tt})(\hat{t}_c)} \right) r \geq \mathfrak{a} r^T Im(\hat{M})(\hat{t}_c) r, \tag{A.1}$$

for some constant $\mathfrak{a}$, if there is no grazing ray,

$$|\hat{p}_d(\hat{t}_c)| \geq \mathfrak{b} ||\hat{p}(\hat{t}_c)||, \tag{A.2}$$

where $\mathfrak{b}$ is an universal lower bound.

By equation (3.60) and (3.61), we have

$$Im(\hat{\tau}_{tx})(\hat{t}_c; y, p) = \mp v(\hat{y}(\hat{t}_c)) Im(\hat{M})(\hat{t}_c) \frac{\hat{p}(\hat{t}_c)}{||\hat{p}(\hat{t}_c)||}, \tag{A.3}$$

and

$$Im(\hat{\tau}_{tt})(\hat{t}_c; y, p) = v^2(\hat{y}(\hat{t}_c))\frac{\hat{p}^T(\hat{t}_c)}{||\hat{p}(\hat{t}_c)||}Im(\hat{M})(\hat{t}_c)\frac{\hat{p}(\hat{t}_c)}{||\hat{p}(\hat{t}_c)||}. \tag{A.4}$$

From now on, we let all functions be defined at $t = \hat{t}_c$ without writing it out explicitly. Then,

$$r^T\frac{Im(\hat{\tau}_{tx})Im(\hat{\tau}_{tx})^T}{Im(\hat{\tau}_{tt})}r = r^T\frac{Im(\hat{M})\hat{p}\hat{p}^T Im(\hat{M})}{\hat{p}^T Im(\hat{M})\hat{p}}r. \tag{A.5}$$

We introduce the following notation,

$$p' = \frac{\hat{p}}{\sqrt{\hat{p}^T Im(\hat{M})\hat{p}}}. \tag{A.6}$$

Then equation (A.1) can be translated to the following optimization problem. We can instead prove that the optimal value of the following optimization problem with fixed p' is $\mathfrak{a}$,

$$\begin{aligned} \min_r \quad & 1 - \left(r^T Im(\hat{M})p'\right)^2, \\ \text{subject to} \quad & r^T Im(\hat{M})r = 1, \quad r^T e_d = 0, \end{aligned} \tag{A.7}$$

where $e_d = (0, \cdots, 0, 1)^T$. It is equivalent to,

$$\begin{aligned} \min_r \quad & 1 - r^T Im(\hat{M})p', \\ \text{subject to} \quad & r^T Im(\hat{M})r = 1, \quad r^T e_d = 0, \end{aligned} \tag{A.8}$$

since the optimization problem (A.7) is the square term of problem (A.8). Using

Lagrange multipliers,

$$L(r;\lambda_1,\lambda_2) = 1 - r^T Im(\hat{M})p^{'} + \lambda_1 r^T Im(\hat{M})r + \lambda_2 r^T e_d. \tag{A.9}$$

Differentiate L with respect to r,

$$-Im(\hat{M})p^{'} + 2\lambda_1 Im(\hat{M})r + \lambda_2 e_d = 0, \tag{A.10}$$

then

$$2\lambda_1 r = -\lambda_2 Im(\hat{M})^{-1} e_d + p^{'}. \tag{A.11}$$

By the second restriction $r^T e_d = 0$,

$$\lambda_2 = \frac{e_d^T p^{'}}{e_d^T Im(\hat{M})^{-1} e_d^T} \tag{A.12}$$

We denote $A = \sqrt{\hat{p}^T Im(\hat{M})\hat{p}}$ and $D = e_d^T Im(\hat{M})^{-1} e_d^T$. Therefore,

$$\lambda_2 = \frac{e_d^T p^{'}}{AD}. \tag{A.13}$$

and the optimizer $r_\star$ satisfies,

$$r_\star \parallel r_0, \tag{A.14}$$

where r_0 is defined as,

$$r_0 \equiv p^{'} - \frac{e_d^T p^{'}}{AD} Im(\hat{M})^{-1} e_d. \tag{A.15}$$

By restriction $r^T Im(\hat{M}) r = 1$, we have,

$$r_\star = \frac{r_0}{\sqrt{r_0^T Im(\hat{M}) r_0}}, \tag{A.16}$$

Equivalently, $2\lambda_1 = \sqrt{r_0^T Im(\hat{M}) r_0}$. Equation (A.7) then becomes,

$$\min_r 1 - \left(r^T Im(\hat{M}) p^{'}\right)^2 = 1 - \left(\frac{r_0^T}{\sqrt{r_0^T Im(\hat{M}) r_0}} Im(\hat{M}) p^{'}\right)^2. \tag{A.17}$$

The first term needed to be evaluated above is,

$$\begin{aligned} r_0^T Im(\hat{M}) p^{'} &= (p^{'} - \frac{e_d^T p^{'}}{AD} Im(\hat{M})^{-1} e_d)^T Im(\hat{M}) p^{'} \\ &= (p^{'})^T Im(\hat{M}) p^{'} - \frac{(e_d^T p^{'})^2}{A^2 D}. \end{aligned}$$

The second term is

$$\begin{aligned} r_0^T Im(\hat{M}) r_0 &= (p^{'} - \frac{e_d^T p^{'}}{AD} Im(\hat{M})^{-1} e_d)^T Im(\hat{M}) (p^{'} - \frac{e_d^T p^{'}}{AD} Im(\hat{M})^{-1} e_d) \\ &= (p^{'})^T Im(\hat{M}) p^{'} + \frac{(e_d^T p^{'})^2}{(AD)^2} e_d^T Im(\hat{M})^{-1} e_d - \frac{2 e_d^T p^{'}}{AD} e_d^T p^{'} \\ &= (p^{'})^T Im(\hat{M}) p^{'} + \frac{(e_d^T p^{'})^2}{(AD)^2} D - \frac{2 e_d^T p^{'}}{AD} \frac{e_d^T p^{'}}{A} \\ &= (p^{'})^T Im(\hat{M}) p^{'} + \frac{(e_d^T p^{'})^2}{A^2 D} - \frac{2 (e_d^T p^{'})^2}{A^2 D} \\ &= (p^{'})^T Im(\hat{M}) p^{'} - \frac{(e_d^T p^{'})^2}{A^2 D}. \end{aligned} \tag{A.18}$$

$$1-\left(\frac{r_0^T}{\sqrt{r_0^T Im(\hat{M})r_0}} Im(\hat{M})p^{'}\right)^2 = 1-\left(\frac{r_0^T Im(\hat{M})p^{'}}{\sqrt{r_0^T Im(\hat{M})r_0}}\right)^2$$

$$= 1-\frac{\left((p^{'})^T Im(\hat{M})p^{'} - \frac{(e_d^T p^{'})^2}{A^2 D}\right)^2}{(p^{'})^T Im(\hat{M})p^{'} - \frac{(e_d^T p^{'})^2}{A^2 D}}$$

$$= \frac{(e_d^T p^{'})^2}{A^2 D}. \tag{A.19}$$

To summarize,

$$r^T \left(Im(\hat{M}) - (r-\hat{r}_0)^T \frac{Im(\hat{\tau}_{tx}) Im(\hat{\tau}_{tx})^T}{Im(\hat{\tau}_{tt})}\right) r \geq \mathfrak{a} r^T Im(\hat{M})(\hat{t}_c) r \tag{A.20}$$

and

$$\mathfrak{a} = \frac{\mathfrak{b}^2}{A^2 D} = \frac{\mathfrak{b}^2}{(\hat{p}^T Im(\hat{M})\hat{p})(e_d^T (Im(\hat{M}))^{-1} e_d)}. \tag{A.21}$$

Proof of Proposition 3.3.5

Proposition A.0.1. *Consider two scattering beams $(\hat{y}(t), \hat{p}(t), \hat{M}(t), \hat{A}(t))$ and $(\hat{x}(t), \hat{\xi}(t), \hat{N}(t), \hat{C}(t))$, and assume that there exists significant interaction effects between these two beams. There exists two constants C_1^* and C_2^* related to the background velocity, such that*

$$\frac{d||\hat{M}(t) - \hat{N}(t)||}{dt} \leq C_1^* \sqrt{||p||} + C_2^* ||\hat{M}(t) - \hat{N}(t)||. \tag{A.22}$$

where $||\hat{M}(t) - \hat{N}(t)||$ is defined as the matrix norm induced by the vector 2-norm.

Proof. Both $\hat{M}(t)$ and $\hat{N}(t)$ satisfy

$$\begin{aligned}\frac{d\hat{M}}{dt} &= -G_{xx}(\hat{y}(t),\hat{p}(t)) - \hat{M}G_{xp}(\hat{y}(t),\hat{p}(t)) - G_{xp}^T(\hat{y}(t),\hat{p}(t))\hat{M} - \hat{M}G_{pp}(\hat{y}(t),\hat{p}(t))\hat{M},\\ \frac{d\hat{N}}{dt} &= -G_{xx}(\hat{x}(t),\hat{\xi}(t)) - \hat{N}G_{xp}(\hat{x}(t),\hat{\xi}(t)) - G_{xp}^T(\hat{x}(t),\hat{\xi}(t))\hat{N} - \hat{N}G_{pp}(\hat{x}(t),\hat{\xi}(t))\hat{N}.\end{aligned} \tag{A.23}$$

Take the difference between above equations, we have

$$\begin{aligned}&-\frac{d(\hat{M}(t)-\hat{N}(t))}{dt} = \\ &\Big(G_{xx}(\hat{y}(t),\hat{p}(t)) - G_{xx}(\hat{x}(t),\hat{\xi}(t))\Big) + 2\hat{N}(t)\left(G_{xp}(\hat{y}(t),\hat{p}(t)) - G_{xp}(\hat{x}(t),\hat{\xi}(t))\right)\\ &+ \hat{N}(t)\left(G_{pp}(\hat{y}(t),\hat{p}(t)) - G_{pp}(\hat{x}(t),\hat{\xi}(t))\right)\hat{N}(t)\\ &+ 2(\hat{M}(t)-\hat{N}(t))G_{xp}(\hat{y}(t),\hat{p}(t)) + (\hat{M}(t)-\hat{N}(t))G_{pp}(\hat{y}(t),\hat{p}(t))(\hat{M}(t)+\hat{N}(t)).\end{aligned} \tag{A.24}$$

The first term in equation (A.24),

$$\begin{aligned}||G_{xx}(\hat{y}(t),\hat{p}(t)) - G_{xx}(\hat{x}(t),\hat{\xi}(t))|| &= \Big|\Big|\nabla\nabla v(\hat{y}(t))||\hat{p}(t)|| - \nabla\nabla v(\hat{x}(t))||\hat{\xi}(t)||\Big|\Big|\\ &\le C_3(||\hat{y}(t)-\hat{x}(t)||)||\hat{p}(t)|| + C_2(||\hat{p}(t)-\hat{\xi}(t)||),\end{aligned} \tag{A.25}$$

where C_3 is the maximum value of the third order derivative of the velocity v and C_2 maximum value of the second order derivative of the velocity v. According to Lemma 3.3.2, we have

$$||G_{xx}(\hat{y}(t),\hat{p}(t)) - G_{xx}(\hat{x}(t),\hat{\xi}(t))|| \sim O(\sqrt{||\hat{p}(t)||}). \tag{A.26}$$

The second term in equation (A.24),

$$
\begin{aligned}
&2\Big|\Big|\hat{N}(t)\left(G_{xp}(\hat{y}(t),\hat{p}(t)) - G_{xp}(\hat{x}(t),\hat{\xi}(t))\right)\Big|\Big| \\
&= 2\Big|\Big|\hat{N}(t)\left(\nabla v(\hat{y}(t))\left(\frac{\hat{p}(t)}{||\hat{p}(t)||}\right)^T - \nabla v(\hat{x}(t))\left(\frac{\hat{\xi}(t)}{||\hat{\xi}(t)||}\right)^T\right)\Big|\Big| \\
&\leq 2||\hat{N}(t)||\left(C_2||\hat{y}(t)-\hat{x}(t)|| + C_1\left(\frac{||\hat{\Xi}(t)||}{||\hat{p}(t)||}\right)\right), \qquad \text{(A.27)}
\end{aligned}
$$

where C_1 is the largest value of $||\nabla v||$. Therefore,

$$
2\Big|\Big|\hat{N}(t)\left(G_{xp}(\hat{y}(t),\hat{p}(t)) - G_{xp}(\hat{x}(t),\hat{\xi}(t))\right)\Big|\Big| \sim O(\sqrt{||\hat{p}(t)||}). \qquad \text{(A.28)}
$$

The third term in equation (A.24),

$$
\begin{aligned}
&\Big|\Big|\hat{N}(t)\left(G_{pp}(\hat{y}(t),\hat{p}(t)) - G_{pp}(\hat{x}(t),\hat{\xi}(t))\right)\hat{N}(t)\Big|\Big| \leq \Big|\Big|\left(\frac{v(\hat{y}(t))}{||\hat{p}(t)||} - \frac{v(\hat{x}(t))}{||\hat{\xi}(t)||}\right)\Big|\Big|\Big|\Big|\hat{N}(t)\hat{N}(t)\Big|\Big| \\
&+ \Big|\Big|\hat{N}(t)\left(\frac{v(\hat{y}(t))}{||\hat{p}(t)||}\left(\frac{\hat{p}(t)}{||\hat{p}(t)||}\right)\left(\frac{\hat{p}(t)}{||\hat{p}(t)||}\right)^T - \frac{v(\hat{x}(t))}{||\hat{\xi}(t)||}\left(\frac{\hat{\xi}(t)}{||\hat{\xi}(t)||}\right)\left(\frac{\hat{\xi}(t)}{||\hat{\xi}(t)||}\right)^T\right)\hat{N}(t)\Big|\Big|.
\end{aligned}
\qquad \text{(A.29)}
$$

First, we have

$$
\begin{aligned}
&\Big|\Big|\frac{v(\hat{y}(t))}{||\hat{p}(t)||} - \frac{v(\hat{x}(t))}{||\hat{\xi}(t)||}\Big|\Big|\Big|\Big|\hat{N}(t)\hat{N}(t)\Big|\Big| \leq \\
&\left(C_1\frac{||\hat{y}(t)-\hat{x}(t)||}{||\hat{p}(t)||} + v(\hat{y}(t))\frac{||\hat{p}(t)-\hat{\xi}(t)||}{||\hat{p}(t)||^2}\right)||\hat{N}(t)\hat{N}(t)|| \sim O(\sqrt{||\hat{p}(t)||}). \qquad \text{(A.30)}
\end{aligned}
$$

Second,

$$\Big|\Big|\frac{v(\hat{y}(t))}{||\hat{p}(t)||}\left(\frac{\hat{p}(t)}{||\hat{p}(t)||}\right)\left(\frac{\hat{p}(t)}{||\hat{p}(t)||}\right)^T - \frac{v(\hat{x}(t))}{||\hat{\xi}(t)||}\left(\frac{\hat{\xi}(t)}{||\hat{\xi}(t)||}\right)\left(\frac{\hat{\xi}(t)}{||\hat{\xi}(t)||}\right)^T\Big|\Big| \leq$$
$$C_1\Big|\Big|\frac{||\hat{y}(t)-\hat{x}(t)||}{||\hat{p}(t)||}\Big|\Big| + C_1\Big|\Big|\frac{v(\hat{y}(t))(1-\hat{\kappa}(t))||\hat{p}(t)||^2}{||\hat{p}(t)||^3}\Big|\Big| \leq O(||\hat{p}(t)||^{-3/2}). \tag{A.31}$$

To summarize,

$$\Big|\Big|\hat{N}(t)\left(G_{pp}(\hat{y}(t),\hat{p}(t)) - G_{pp}(\hat{x}(t),\hat{\xi}(t))\right)\hat{N}(t)\Big|\Big| \sim O(\sqrt{||\hat{p}(t)||}). \tag{A.32}$$

Insert these asymptotic analysis equations (A.26), (A.28), (A.32) into equation (A.24), we have

$$\begin{aligned}
&-\frac{d(\hat{M}(t)-\hat{N}(t))}{dt} = O(\sqrt{||\hat{p}(t)||}) + 2(\hat{M}(t)-\hat{N}(t))G_{xp}(\hat{y}(t),\hat{p}(t))\\
&+(\hat{M}(t)-\hat{N}(t))G_{pp}(\hat{y}(t),\hat{p}(t))(\hat{M}(t)+\hat{N}(t))\\
&= O(\sqrt{||\hat{p}(t)||}) + (\hat{M}(t)-\hat{N}(t))\left(2G_{xp}(\hat{y}(t),\hat{p}(t)) + G_{pp}(\hat{y}(t),\hat{p}(t))(\hat{M}(t)+\hat{N}(t))\right).
\end{aligned} \tag{A.33}$$

The coefficient matrix in front of $\hat{M}(t)-\hat{N}(t)$ satisfies,

$$\begin{aligned}
||2G_{xp}(\hat{y}(t),\hat{p}(t)) + G_{pp}(\hat{y}(t),\hat{p}(t))(\hat{M}(t)+\hat{N}(t))|| &\leq 2||\nabla v(\hat{y}(t))|| + 2\frac{||\hat{M}(t)+\hat{N}(t)||}{||\hat{p}(t)||}\\
&\leq O(1).
\end{aligned} \tag{A.34}$$

Then, $\hat{M}(t)-\hat{N}(t)$ satisfies,

$$\frac{d||\hat{M}(t)-\hat{N}(t)||}{dt} \leq C_1^*\sqrt{||p||} + C_2^*||\hat{M}(t)-\hat{N}(t)||, \quad ||\hat{M}(0)-\hat{N}(0)|| = 0. \tag{A.35}$$

$\square$

Proof of Lemma 3.3.5

Lemma A.0.2. *Consider two scattering beams* $(\hat{y}(t), \hat{p}(t), \hat{M}(t), \hat{A}(t))$ *and* $(\hat{x}(t), \hat{\xi}(t), \hat{N}(t), \hat{C}(t))$*, and there exists significant interaction effects between these two beams. Suppose the function* $g(t)$ *is*

$$g(t) = \hat{p}(t) \cdot (\hat{y}(t) - \hat{x}(t)), \tag{A.36}$$

then we have

$$g(t) = g(0) + O(1), \tag{A.37}$$

and

$$g^{'}(t) = v(\hat{x}(t)) \left(\frac{1}{2} \frac{||\hat{\Xi}(t)||^2}{\hat{\kappa}(t)^2 ||\hat{p}(t)||} \right). \tag{A.38}$$

Proof. We start with calculating the derivative $\frac{dg(t)}{dt}$ and assume the Hamiltonian to be positive, i.e. $G(x,p) = v(x)||p||$. Negative branch will be the same.

$$\begin{aligned} \frac{dg}{dt} &= \frac{d\hat{p}(t)}{dt} \cdot (\hat{y}(t) - \hat{x}(t)) + \hat{p}(t) \cdot \left(\frac{d\hat{y}(t)}{dt} - \frac{d\hat{x}(t)}{dt} \right) \\ &= -\nabla v(\hat{y}(t)) \cdot (\hat{y}(t) - \hat{x}(t)) ||\hat{p}(t)|| + \hat{p}(t) \cdot \left(v(\hat{y}(t)) \frac{\hat{p}(t)}{||\hat{p}(t)||} - v(\hat{x}(t)) \frac{\hat{\xi}(t)}{||\hat{\xi}(t)||} \right). \end{aligned} \tag{A.39}$$

Use the Taylor expansion of the velocity v at the point $\hat{y}(t)$,

$$
\begin{aligned}
\frac{dg}{dt} &= ||\hat{p}(t)|| \left(v(\hat{x}(t)) - v(\hat{y}(t))\right) + ||\hat{p}(t)|| v(\hat{y}(t)) - v(\hat{x}(t)) \frac{\hat{p}(t) \cdot \hat{\xi}(t)}{||\hat{\xi}(t)||} + O(||\hat{x}(t) - \hat{y}(t)||^2) \\
&= ||\hat{p}(t)|| v(\hat{x}(t)) - v(\hat{x}(t)) \frac{\hat{p}(t) \cdot \hat{\xi}(t)}{||\hat{\xi}(t)||} + O(||\hat{x}(t) - \hat{y}(t)||^2) \\
&= v(\hat{x}(t)) \left(||\hat{p}(t)|| - \frac{\hat{p}(t) \cdot \hat{\xi}(t)}{||\hat{\xi}(t)||} \right) + O(||\hat{x}(t) - \hat{y}(t)||^2).
\end{aligned}
$$

Substitute decomposition (3.56) into the fraction term $\frac{\hat{p}(t) \cdot \hat{\xi}(t)}{||\hat{\xi}(t)||}$. Since $\hat{\Xi}(t)$ is orthogonal to $\hat{p}(t)$,

$$
\begin{aligned}
\frac{\hat{p}(t) \cdot \hat{\xi}(t)}{||\hat{\xi}(t)||} &= \frac{\hat{\kappa}(t) ||\hat{p}(t)||^2}{\sqrt{\hat{\kappa}(t)^2 ||\hat{p}(t)||^2 + ||\hat{\Xi}(t)||^2}} \\
&= \frac{\hat{\kappa}(t) ||\hat{p}(t)||^2}{\hat{\kappa}(t) ||\hat{p}(t)||} \frac{1}{1 + \frac{1}{2} \frac{||\hat{\Xi}(t)||^2}{||\hat{\kappa}(t) \hat{p}(t)||^2}} \\
&= ||\hat{p}(t)|| - \frac{1}{2} \frac{||\hat{\Xi}(t)||^2}{|\hat{\kappa}(t)|^2 ||\hat{p}(t)||}.
\end{aligned}
$$

Here we use the Taylor expansion of the square root function and Geometric series to approximate. Then

$$
\begin{aligned}
\frac{dg}{dt} &= v(\hat{x}(t)) \left(||\hat{p}(t)|| - \frac{\hat{\kappa}(t) ||\hat{p}(t)||^2}{\hat{\kappa}(t) ||\hat{p}(t)||} \right) + \frac{v(\hat{x}(t))}{2} \frac{||\hat{\Xi}(t)||^2}{|\hat{\kappa}(t)|^2 ||\hat{p}(t)||} \\
&= O(1).
\end{aligned} \tag{A.40}
$$

Naturally, after finite time,

$$
g(t) = g(0) + O(1). \tag{A.41}
$$

□

Proof of Lemma 3.3.6

Lemma A.0.3. *Consider two scattering beams* $(\hat{y}(t), \hat{p}(t), \hat{M}(t), \hat{A}(t))$ *and* $(\hat{x}(t), \hat{\xi}(t), \hat{N}(t), \hat{C}(t))$*, and there exists significant interaction effects between these two beams. Suppose the pure imaginary matrix* $\hat{M}(0)$ *has a symmetric positive definite imaginary part and is the initial condition of the Hessian for the beam, then*

$$(y-x)^T \hat{M}(0)(y-x) = (\hat{y}(t) - \hat{x}(t))^T \hat{M}(t)(\hat{y}(t) - \hat{x}(t)) + O(1). \tag{A.42}$$

Proof. We denote the function $g(t)$ as

$$g(t) = ||\hat{y}(t) - \hat{x}(t)||^2_{\hat{M}(t)}$$

Throughout this proof, we use the Hamiltonian $G(x, p) = v(x)||p||$ and the negative Hamiltonian will be the same. The derivative of g is,

$$\begin{aligned}
&\frac{dg(t)}{dt} = 2(G_p(\hat{y}(t), \hat{p}(t)) - G_p(\hat{x}(t), \hat{\xi}(t)))^T \hat{M}(t)(\hat{y}(t) - \hat{x}(t)) \\
&+ (\hat{y}(t) - \hat{x}(t))^T \frac{d\hat{M}(t)}{dt}(\hat{y}(t) - \hat{x}(t)) \\
&= 2\left(G_{px}(\hat{y}(t), \hat{p}(t))(\hat{y}(t) - \hat{x}(t)) + G_{pp}(\hat{y}(t), \hat{p}(t))(\hat{p}(t) - \hat{\xi}(t))\right)^T \hat{M}(t)(\hat{y}(t) - \hat{x}(t)) \\
&+ (\hat{y}(t) - \hat{x}(t))^T \frac{d\hat{M}(t)}{dt}(\hat{y}(t) - \hat{x}(t)).
\end{aligned} \tag{A.43}$$

From Riccati equation (2.11), we have

$$\frac{dM}{dt} = -G_{xx} - MG_{xp} - G_{xp}^T M - MG_{pp}M.$$

Insert the above equation into equation (A.43) and abbreviate $G = G(\hat{y}(t), \hat{p}(t))$, if there's no variable specified,

$$\frac{dg(t)}{dt} = -(\hat{y}(t) - \hat{x}(t))^T G_{xx}(\hat{y}(t) - \hat{x}(t)) - 2(\hat{y}(t) - \hat{x}(t))^T \hat{M}(t) G_{xp}(\hat{y}(t) - \hat{x}(t))$$
$$+ 2(\hat{y}(t) - \hat{x}(t))^T \hat{M}(t) G_{px}(\hat{y}(t) - \hat{x}(t))$$
$$+ 2(\hat{p}(t) - \hat{\xi}(t))^T G_{pp} \hat{M}(t)(\hat{y}(t) - \hat{x}(t)) - \left(\hat{M}(t)(\hat{y}(t) - \hat{x}(t))\right)^T G_{pp} \left(\hat{M}(t)(\hat{y}(t) - \hat{x}(t))\right)$$
$$= -(\hat{y}(t) - \hat{x}(t))^T G_{xx}(\hat{y}(t) - \hat{x}(t)) + 2(\hat{p}(t) - \hat{\xi}(t))^T G_{pp} \hat{M}(t)(\hat{y}(t) - \hat{x}(t))$$
$$- \left(\hat{M}(t)(\hat{y}(t) - \hat{x}(t))\right)^T G_{pp} \left(\hat{M}(t)(\hat{y}(t) - \hat{x}(t))\right)$$
$$= -(\hat{y}(t) - \hat{x}(t))^T G_{xx}(\hat{y}(t) - \hat{x}(t)) - ||\hat{M}(t)(\hat{y}(t) - \hat{x}(t)) - (\hat{p}(t) - \hat{\xi}(t))||^2_{G_{pp}}$$
$$+ ||\hat{p}(t) - \hat{\xi}(t)||^2_{G_{pp}}. \tag{A.44}$$

The first term in equation (A.44) satisfies,

$$(\hat{y}(t) - \hat{x}(t))^T (G_{xx}(\hat{y}(t), \hat{p}(t)))(\hat{y}(t) - \hat{x}(t)) \sim O(1). \tag{A.45}$$

First, $G_{xx}(\hat{y}(t), \hat{p}(t)) \sim O(||\hat{p}(t)||)$, since v is smooth. On the other hand, $\hat{y}(t) - \hat{x}(t)$'s order is $O(1/\sqrt{||\hat{p}(t)||})$ by Lemma 3.3.2.

The second term in equation (A.44) satisfies,

$$||\hat{M}(t)(\hat{y}(t) - \hat{x}(t)) - (\hat{p}(t) - \hat{\xi}(t))||^2_{G_{pp}(\hat{y}(t),\hat{p}(t))} \sim O(||\hat{p}(t)||^{1/2} ||\hat{p}(t)||^{-1} ||\hat{p}(t)||^{1/2}).$$
(A.46)

The term $G_{pp}(\hat{y}(t), \hat{p}(t))$ is of the order $O(1/||\hat{p}(t)||)$ implied by G_{pp}'s expression. Again, Lemma 3.3.2 demonstrates that $\hat{M}(t)(\hat{y}(t) - \hat{x}(t)) - (\hat{p}(t) - \hat{\xi}(t)) \sim O(\sqrt{||\hat{p}(t)||})$.

The third term in equation (A.44) satisfies,

$$||\hat{p}(t) - \hat{\xi}(t)||^2_{G_{pp}(\hat{y}(t),\hat{p}(t))} \sim O(||\hat{p}(t)||^{1/2}||\hat{p}(t)||^{-1}||\hat{p}(t)||^{1/2}) \sim O(1). \tag{A.47}$$

It can be justified by combining G_{pp}'s expression and the fact that $||\hat{p}(t) - \hat{\xi}(t)|| \sim O(\sqrt{||\hat{p}(t)||})$. □

Proof of Proposition 3.3.3

Proposition A.0.2.

$$\begin{aligned}
&i\hat{\gamma}(\hat{t}_0(x,\xi);x,\xi)|\omega - \hat{\tau}_t(\hat{t}_0(x,\xi);x,\xi) - \hat{\vartheta}(\hat{t}_0(x,\xi);x,\xi)^T(r - \hat{x}(\hat{t}_0(x,\xi)))|^2 \\
&\approx i\hat{\gamma}(\hat{t}_0(y,p);x,\xi)|\omega - \hat{\tau}_t(\hat{t}_0(y,p);x,\xi) - \left(\hat{\vartheta}(\hat{t}_0(y,p);x,\xi)\right)^T(r - \hat{x}(\hat{t}_0(y,p)))|^2 \\
&+ O\left(\frac{1}{||p||}\right).
\end{aligned} \tag{A.48}$$

Proof. The coefficient $\hat{\gamma}$ satisfies,

$$\hat{\gamma}(\hat{t}_0(x,\xi);x,\xi) = -\frac{1}{2\hat{\tau}_{tt}(\hat{t}_0(x,\xi);x,\xi)}.$$

By Assumption 3.2.3, it is safe to say that the background velocity v around $\hat{x}(\hat{t}_0(y,p))$ is a constant function. Consequently, we have Hamiltonian satisfying $G_x = 0$, and

$$\begin{aligned}
&\hat{G}_p(\hat{x}(\hat{t}_0(x,\xi);x,\xi),\hat{\xi}(\hat{t}_0(x,\xi);x,\xi)) = \pm v(\hat{x}(\hat{t}_0(y,p)))\frac{\hat{\xi}(\hat{t}_0(y,p)))}{||\hat{\xi}(\hat{t}_0(y,p))||} \\
&= \hat{G}_p(\hat{x}(\hat{t}_0(y,p);x,\xi),\hat{\xi}(\hat{t}_0(y,p);x,\xi)),
\end{aligned} \tag{A.49}$$

since ray direction does not change in the constant slowness. Then

$$\begin{aligned}
\hat{\tau}_{tt}(\hat{t}_0(x,\xi);x,\xi) &= \hat{G}_p^T(\hat{x}(\hat{t}_0(x,\xi)),\hat{\xi}(\hat{t}_0(x,\xi)))\hat{M}(\hat{t}_0(x,\xi);x,\xi)\hat{G}_p(\hat{x}(\hat{t}_0(x,\xi)),\hat{\xi}(\hat{t}_0(x,\xi))) \\
&= \hat{G}_p^T(\hat{x}(\hat{t}_0(y,p)),\hat{\xi}(\hat{t}_0(y,p)))\left(\hat{M}(\hat{t}_0(y,p);x,\xi)+\frac{d\hat{M}}{dt}(\hat{t}_0(x,\xi)-\hat{t}_0(y,p))\right) \\
&\quad \hat{G}_p(\hat{x}(\hat{t}_0(y,p)),\hat{\xi}(\hat{t}_0(y,p))) \\
&= \hat{\tau}_{tt}(\hat{t}_0(y,p);x,\xi)+O(\sqrt{||p||}), \qquad (A.50)
\end{aligned}$$

since we have $\frac{d\hat{M}}{dt} \sim O(||p||)$ by Lemma 3.3.6. On the other hand, $|\hat{t}_0(y,p)-\hat{t}_0(x,\xi)| \sim O(||p||^{-1/2})$.

We then have

$$\begin{aligned}
\hat{\gamma}(\hat{t}_0(x,\xi);x,\xi) &= \frac{1}{2(\hat{\tau}_{tt}(\hat{t}_0(y,p);x,\xi)+O(\sqrt{||p||}))} \\
&= \frac{1}{2(\hat{\tau}_{tt}(\hat{t}_0(y,p);x,\xi)}\left(\frac{1}{1+(\sqrt{||p||})^{-1}}\right) \\
&\approx \frac{1}{2(\hat{\tau}_{tt}(\hat{t}_0(y,p);x,\xi)} = \hat{\gamma}(\hat{t}_0(y,p);x,\xi). \qquad (A.51)
\end{aligned}$$

Inside the quadratic term, we first have an invariant,

$$\omega - \hat{\tau}_t(\hat{t}_0(x,\xi);x,\xi) = \omega - \hat{\tau}_t(\hat{t}_0(y,p);x,\xi), \qquad (A.52)$$

since $\hat{\tau}_t$ will be a constant along the ray.

The next term is $\hat{\vartheta}(\hat{t}_0(x,\xi);x,\xi)$.

$$Re(\hat{\tau}_{tx}(\hat{t}_0(x,\xi);x,\xi)) = -Re(\hat{M})(\hat{t}_0(x,\xi);x,\xi)G_p(\hat{x}(\hat{t}_0(x,\xi);x,\xi),\hat{\xi}(\hat{t}_0(x,\xi);x,\xi))$$

$$= -Re(\hat{M})(\hat{t}_0(y,p);x,\xi)G_p(\hat{x}(\hat{t}_0(y,p);x,\xi),\hat{\xi}(\hat{t}_0(y,p);x,\xi))$$

$$+ \frac{dRe(\hat{M})(\hat{t}_0(y,p);x,\xi)}{dt}(\hat{t}_0(x,\xi)-\hat{t}_0(y,p))G_p(\hat{x}(\hat{t}_0(y,p);x,\xi),\hat{\xi}(\hat{t}_0(y,p);x,\xi))$$

$$= Re(\hat{\tau}_{tx}(\hat{t}_0(y,p);x,\xi)) + O(\sqrt{||p||}). \tag{A.53}$$

We have the similar conclusion for imaginary part,

$$Im(\hat{\tau}_{tx}(\hat{t}_0(x,\xi);x,\xi)) = Im(\hat{\tau}_{tx}(\hat{t}_0(y,p);x,\xi)) + O(\sqrt{||p||}). \tag{A.54}$$

Consequently, by equation (3.75)

$$\hat{\vartheta}(\hat{t}_0(x,\xi);x,\xi)) = \hat{\vartheta}(\hat{t}_0(y,p);x,\xi)) + O(\sqrt{||p||}). \tag{A.55}$$

Next,

$$r - \hat{x}(\hat{t}_0(x,\xi)) = r - \hat{x}(\hat{t}_0(y,p)) + \hat{x}(\hat{t}_0(y,p)) - \hat{x}(\hat{t}_0(x,\xi))$$

$$= r - \hat{x}(\hat{t}_0(y,p)) + G_p^{\pm}(\hat{x}(\hat{t}_0(y,p)),\hat{\xi}(\hat{t}_0(y,p)))(\hat{t}_0(y,p)-\hat{t}_0(x,\xi))$$

$$= r - \hat{x}(\hat{t}_0(y,p)) \pm v(\hat{x}(\hat{t}_0(y,p)))(\hat{t}_0(y,p)-\hat{t}_0(x,\xi))\frac{\hat{\xi}(\hat{t}_0(y,p))}{||\hat{\xi}(\hat{t}_0(y,p))||} \tag{A.56}$$

$$= r - \hat{x}(\hat{t}_0(y,p)) + O(1/\sqrt{||p||}). \tag{A.57}$$

$G^{\pm}$ represents the different Hamiltonian by its sign. Therefore,

$$
\begin{aligned}
&i\hat{\gamma}(\hat{t}_0(x,\xi);x,\xi)|\omega - \hat{\tau}_t(\hat{t}_0(x,\xi);x,\xi) - \hat{\vartheta}(\hat{t}_0(x,\xi);x,\xi)^T(r - \hat{x}(\hat{t}_0(x,\xi)))|^2 = i\hat{\gamma}(\hat{t}_0(y,p);x,\xi) \\
&|\omega - \hat{\tau}_t(\hat{t}_0(y,p);x,\xi) - \left(\hat{\vartheta}(\hat{t}_0(y,p);x,\xi) + O(\sqrt{||p||})\right)^T (r - \hat{x}(\hat{t}_0(x,\xi))|^2 + O\left(\frac{1}{\sqrt{||p||}}\right) \\
&\approx i\hat{\gamma}(\hat{t}_0(y,p);x,\xi)|\omega - \hat{\tau}_t(\hat{t}_0(y,p);x,\xi) - \left(\hat{\vartheta}(\hat{t}_0(y,p);x,\xi)\right)^T (r - \hat{x}(\hat{t}_0(y,p)))|^2.
\end{aligned}
\tag{A.58}
$$

□

Proof of Proposition 3.3.6

Proposition A.0.3. *Some real-valued phase terms, $\hat{\varrho}(r,\hat{t}_c;y,p) - \hat{\theta}(r,\hat{t}_c;y,p)$ and $\hat{\varrho}(r,\hat{t}_c;x,\xi) - \hat{\theta}(r,\hat{t}_c;x,\xi)$, can be ignored since they are constant order terms with respect to the large wavenumber $||\xi_{l,i}|| = ||p+q||$.*

Proof. **The first term:** $(\hat{F} - \omega)\hat{Q}$

$$
\begin{aligned}
&(\hat{\tau}_t(\hat{t}_c;y,p) - \omega + Re(\hat{\tau}_{tx})(\hat{t}_c;y,p)) \left(\frac{Im(\hat{\tau}_{tx}(\hat{t}_c;y,p))^T(r - \hat{y}(\hat{t}_c))}{Im(\hat{\tau}_{tt}(\hat{t}_c;y,p))}\right) \\
&- (\hat{\tau}_t(\hat{t}_c;x,\xi) - \omega + Re(\hat{\tau}_{tx})(\hat{t}_c;x,\xi)) \left(\frac{Im(\hat{\tau}_{tx}(\hat{t}_c;t,\xi))^T(r - \hat{x}(\hat{t}_c))}{Im(\hat{\tau}_{tt}(\hat{t}_c;x,\xi))}\right) \\
&\approx (\hat{\tau}_t(\hat{t}_c;x,\xi) - \omega)(\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c)) + (\hat{\tau}_t(\hat{t}_c;y,p) - \hat{\tau}_t(\hat{t}_c;x,\xi))(r - \hat{y}(\hat{t}_c)) + O(1).
\end{aligned}
$$

By Corollary 3.3.1, we notice the scale of $|\hat{\tau}_t(\hat{t}_c;x,\xi)$ is controlled by $Im(\hat{\gamma}) \sim O(\frac{1}{||p||})$.

Therefore, by Lemma 3.3.2, $\Big|\hat{\tau}_t(\hat{t}_c;x,\xi)-\omega\Big|||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||\sim O(1)$.

$$
\begin{aligned}
|\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\tau}_t(\hat{t}_c;x,\xi)| &= \Big|v(\hat{y}(\hat{t}_c))||\hat{p}(\hat{t}_c)||-v(\hat{x}(\hat{t}_c))||\hat{\xi}(\hat{t}_c)||\Big| \\
&\leq |\nabla v(\hat{y}(\hat{y}(\hat{t}_c))|||(\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||||\hat{p}(\hat{y}(\hat{t}_c))|| \\
&+ v(\hat{y}(\hat{t}_c))(||\hat{p}(\hat{t}_c)||-||\hat{\xi}(\hat{t}_c)||).
\end{aligned}
$$

By Lemma 3.3.2, $|\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\tau}_t(\hat{t}_c;x,\xi)|\sim O(\sqrt{||p||})$, and $||r-\hat{y}(\hat{t}_c)||$ is controlled by Hessian $\hat{M}(\hat{t}_c)$. So $(\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\tau}_t(\hat{t}_c;x,\xi))(r-\hat{y}(\hat{t}_c))\sim O(1)$.

The second term: $\frac{1}{2}Re(\hat{\tau}_{tt})\hat{Q}^2$

$$
\frac{1}{2}Re(\hat{\tau}_{tt})\hat{Q}^2\approx\frac{1}{2}Re(\hat{\tau}_{tt})(r-\hat{y}(\hat{t}_c))^2\sim O(1), \tag{A.59}
$$

by Lemma 3.3.2. And the same analysis can be applied to the term associated with the beam (x,ξ). $\square$

Proof of Proposition 3.3.7

Proposition A.0.4. *For the first two terms in equation (3.107), their exponents satisfy,*

$$
\begin{aligned}
&-Im(\hat{\beta})|\omega-\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\zeta}^T(r-\hat{y}(\hat{t}_c))|^2-Im(\hat{\gamma})|\omega-\hat{\tau}_t(\hat{t}_c;x,\xi)-\hat{\vartheta}^T(r-\hat{x}(\hat{t}_c))|^2 \\
&=-||r-\frac{\hat{x}(\hat{t}_c)+\hat{y}(\hat{t}_c)}{2}||^2_{2Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T}-\frac{Im(\hat{\beta})}{2}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{\hat{\zeta}\hat{\zeta}^T} \\
&-Im(\hat{\beta})|\omega-\hat{\tau}_t(\hat{t}_c;y,p)|^2-Im(\hat{\gamma})|\omega-\hat{\tau}_t(\hat{t}_c;x,\xi)|^2+O\left(\frac{1}{\sqrt{||p||}}\right),
\end{aligned} \tag{A.60}
$$

where $\hat{\beta}$, $\hat{\gamma}$, $\hat{\zeta}$ and $\hat{\vartheta}$ are all defined at $\hat{t}_c$. Similarly, $\hat{g}$ in equation (3.108) is an $O(1)$ term.

Proof. The quadratic term about r in equation (A.60) is,

$$-r^T \left(Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T + Im(\hat{\gamma})\hat{\vartheta}(\hat{\vartheta})^T \right) r. \tag{A.61}$$

The matrix in the parentheses is positive semi-definite matrix as it is the sum of two positive semi-definite matrices. We apply the eigenvalue decomposition

$$Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T + Im(\hat{\gamma})\hat{\vartheta}(\hat{\vartheta})^T = Q^T \Lambda Q, \tag{A.62}$$

where Λ is a diagonal matrix with non-negative entries and $Q^T Q = I$.

The cross term about r in equation (A.60) is

$$\begin{aligned} & 2Im(\hat{\beta})(\omega - \hat{\tau}_t(\hat{t}_c; y, p))\hat{\zeta}^T r + 2Im(\hat{\beta})\hat{y}(\hat{t}_c)^T(\hat{\zeta}\hat{\zeta}^T)r \\ & + 2Im(\hat{\gamma})(\omega - \hat{\tau}_t(\hat{t}_c; x, \xi))\hat{\vartheta}^T r + 2Im(\hat{\gamma})\hat{x}(\hat{t}_c)^T(\hat{\vartheta}\hat{\vartheta}^T)r \\ & = 2\hat{J}_1^T(\hat{t}_c, \omega; y, p, x, \xi)r + 2\hat{J}_2^T(\hat{t}_c, \omega; y, p, x, \xi)r, \end{aligned} \tag{A.63}$$

where

$$\hat{J}_1(\hat{t}_c, \omega; y, p, x, \xi) = Im(\hat{\beta})(\omega - \hat{\tau}_t(\hat{t}_c; y, p))\hat{\zeta} + Im(\hat{\gamma})(\omega - \hat{\tau}_t(\hat{t}_c; x, \xi))\hat{\vartheta}; \tag{A.64}$$

$$\hat{J}_2(\hat{t}_c, \omega; y, p, x, \xi) = Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T \hat{y}(\hat{t}_c) + Im(\hat{\gamma})\hat{\vartheta}\hat{\vartheta}^T \hat{x}(\hat{t}_c). \tag{A.65}$$

Combine the cross term (A.63) and second order term (A.61),

$$-(Qr)^T\Lambda(Qr) + 2\hat{J}_1^T Q^T Qr + 2\hat{J}_2^T Q^T Qr. \tag{A.66}$$

To make a complete quadratic term, its central point r_c is,

$$\begin{aligned} r_c &= Q^T\Lambda^{-1}Q(\hat{J}_1 + \hat{J}_2) \\ &= Q^T\Lambda^{-1}Q\hat{J}_2 + O\left(\frac{1}{\sqrt{||p||}}\right), \end{aligned} \tag{A.67}$$

since

$$\begin{aligned} ||\hat{J}_1|| &\leq ||Im(\hat{\beta})(\omega - \hat{\tau}_t(\hat{t}_c; y, p))\hat{\zeta}|| + ||Im(\hat{\gamma})(\omega - \hat{\tau}_t(\hat{t}_c; x, \xi))\hat{\vartheta}|| \\ &\leq O\left(\frac{1}{||p||}\right) O\left(\sqrt{||p||}\right) O\left(||p||\right) \leq O\left(\sqrt{||p||}\right), \end{aligned}$$

and

$$\begin{aligned} ||\hat{J}_2|| &\leq ||Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T\hat{y}(\hat{t}_c)|| + ||Im(\hat{\gamma})\hat{\vartheta}\hat{\vartheta}^T\hat{x}(\hat{t}_c)|| \\ &\leq O\left(\frac{1}{||p||}\right) O\left(||p||^2\right) O(1) \leq O(||p||). \end{aligned}$$

Here the inverse matrix is defined as the pseudo-inverse matrix for a rank-deficient matrix, that is $\Lambda_{ii}^{-1} = 0$, if $\Lambda_{ii} = 0$. If the diagonal term in Λ_{ii}^{-1} is zero, then we set its center r_c's i^{th} coordinate to be the same as $\frac{\hat{y}(\hat{t}_c)+\hat{x}(\hat{t}_c)}{2}$'s, since we have $\exp\left(-\frac{1}{4}||2r - \hat{x}(\hat{t}_c) - \hat{y}(\hat{t}_c)||^2_{\hat{\mathcal{M}}(\hat{t}_c)}\right)$ in the expression of $\hat{B}$. Therefore, we carry out calculation by assuming Λ invertible.

$$\begin{aligned}Q^T\Lambda^{-1}Q &= (Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T + Im(\hat{\gamma})\hat{\vartheta}(\hat{\vartheta})^T)^{-1}\\ &= (2Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T + Im(\hat{\gamma})\hat{\vartheta}(\hat{\vartheta})^T - Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T)^{-1}.\end{aligned}$$

To evaluate the order of $Im(\hat{\gamma})\hat{\vartheta}(\hat{\vartheta})^T - Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T$, we first have

$$\begin{aligned}|\hat{\beta}-\hat{\gamma}| &= \frac{1}{2}\Big|\frac{1}{\hat{\tau}_{tt}(\hat{t}_c;y,p)} - \frac{1}{\hat{\tau}_{tt}(\hat{t}_c;x,\xi)}\Big|\\ &\leq \Big|\frac{\hat{\tau}_{tt}(\hat{t}_c;y,p) - \hat{\tau}_{tt}(\hat{t}_c;x,\xi)}{\hat{\tau}_{tt}(\hat{t}_c;y,p)\hat{\tau}_{tt}(\hat{t}_c;x,\xi)}\Big|. \qquad (A.68)\end{aligned}$$

By Corollary 3.3.3, we have $|\hat{\tau}_{tt}(\hat{t}_c;y,p) - \hat{\tau}_{tt}(\hat{t}_c;x,\xi)| \sim O(\sqrt{||p||})$. Consequently,

$$|\hat{\beta}-\hat{\gamma}| \leq O(||p||^{-\frac{3}{2}}). \qquad (A.69)$$

On the other hand, by using Corollary 3.3.3,

$$||\hat{\zeta}-\hat{\vartheta}|| \leq O(\sqrt{||p||}). \qquad (A.70)$$

Then

$$\begin{aligned}||Im(\hat{\gamma})\hat{\vartheta}(\hat{\vartheta})^T - Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T|| &\leq ||Im(\hat{\beta})\left(\hat{\zeta}\hat{\zeta}^T - \hat{\vartheta}\hat{\vartheta}^T\right)|| + ||Im(\hat{\beta}-\hat{\gamma})\hat{\vartheta}\hat{\vartheta}^T||\\ &\leq O(\sqrt{||p||}). \qquad (A.71)\end{aligned}$$

Therefore,

$$Q^T \Lambda^{-1} Q = (2Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T + O(\sqrt{||p||}))^{-1} \approx (2Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T)^{-1}. \tag{A.72}$$

Similarly,

$$\begin{aligned}
\hat{J}_2 &= Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T(\hat{y}(\hat{t}_c) + \hat{x}(\hat{t}_c)) - Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T\hat{x}(\hat{t}_c) + Im(\hat{\gamma})\hat{\vartheta}(\hat{\vartheta})^T\hat{x}(\hat{t}_c) \\
&= Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T(\hat{y}(\hat{t}_c) + \hat{x}(\hat{t}_c)) + O(\sqrt{||p||}). \\
&\approx Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T(\hat{y}(\hat{t}_c) + \hat{x}(\hat{t}_c)). \qquad (A.73)
\end{aligned}$$

The central point r_c becomes,

$$r_c \approx Q^T \Lambda^{-1} Q \hat{J}_2 \approx (2Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T)^{-1} Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T(\hat{y}(\hat{t}_c) + \hat{x}(\hat{t}_c)) = \frac{\hat{x}(\hat{t}_c) + \hat{y}(\hat{t}_c)}{2}. \tag{A.74}$$

Equation (A.60) now becomes,

$$\begin{aligned}
&- Im(\hat{\beta})|\omega - \hat{\tau}_t(\hat{t}_c; y, p) - \hat{\zeta}^T(r - \hat{y}(\hat{t}_c))|^2 - Im(\hat{\gamma})|\omega - \hat{\tau}_t(\hat{t}_c; x, \xi) - (\hat{\vartheta})^T(r - \hat{x}(\hat{t}_c))|^2 \\
&= -Im(\hat{\beta})|\omega - \hat{\tau}_t(\hat{t}_c; y, p)|^2 - Im(\hat{\gamma})|\omega - \hat{\tau}_t(\hat{t}_c; x, \xi)|^2 - ||r - \frac{\hat{x}(\hat{t}_c) + \hat{y}(\hat{t}_c)}{2}||^2_{2Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T} \\
&+ \hat{J}_2^T(2Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T)^{-1}\hat{J}_2 - Im(\hat{\beta})|\hat{\zeta}^T\hat{y}(\hat{t}_c)|^2 - Im(\hat{\gamma})|(\hat{\vartheta})^T\hat{x}(\hat{t}_c)|^2.
\end{aligned}$$

Using equation (A.71),

$$\hat{J}_2^T(2Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T)^{-1}\hat{J}_2 - Im(\hat{\beta})|\hat{\zeta}^T\hat{y}(\hat{t}_c)|^2 - Im(\hat{\gamma})|(\hat{\vartheta})^T\hat{x}(\hat{t}_c)|^2$$

$$\approx ||\hat{y}(\hat{t}_c) + \hat{x}(\hat{t}_c)||^2_{\frac{Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T}{2}} - ||\hat{y}(\hat{t}_c)||^2_{Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T} - ||\hat{x}(\hat{t}_c)||^2_{Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T}$$

$$= -||\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c)||^2_{\frac{Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T}{2}}. \tag{A.75}$$

Equation (A.60) now becomes,

$$- Im(\hat{\beta})|\omega - \hat{\tau}_t(\hat{t}_c; y, p) - \hat{\zeta}^T(r - \hat{y}(\hat{t}_c))|^2 - Im(\hat{\gamma})|\omega - \hat{\tau}_t(\hat{t}_c; x, \xi) - (\hat{\vartheta})^T(r - \hat{x}(\hat{t}_c))|^2$$

$$\approx -Im(\hat{\beta})|\omega - \hat{\tau}_t(\hat{t}_c; y, p)|^2 - Im(\hat{\gamma})|\omega - \hat{\tau}_t(\hat{t}_c; x, \xi)|^2$$

$$- ||r - \frac{\hat{x}(\hat{t}_c) + \hat{y}(\hat{t}_c)}{2}||^2_{2Im(\hat{\beta})\hat{\zeta}\hat{\zeta}^T} - \frac{Im(\hat{\beta})}{2}||\hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c)||^2_{\hat{\zeta}\hat{\zeta}^T}.$$

By applying the similar computation, the exponent of the last term in equation (3.107) denoted as $\hat{g}$ contains two $O(1)$ terms. □

Proof of Proposition 3.3.8

Proposition A.0.5. *The function $\hat{\phi}_1$ satisfies*

$$\hat{\phi}_1(t; x, \xi, y, p) = \frac{1}{2}||\hat{p}(t) - \hat{\xi}(t)||^2_{G_{pp}} + O\left(\frac{1}{\sqrt{||p||}}\right). \tag{A.76}$$

Proof. Using Lemma 3.3.5 for $\hat{\phi}_1(t)$,

$$\hat{\phi}_1(t; x, \xi, y, p) = v(\hat{x}(t))\left(\frac{1}{2}\frac{||\hat{\Xi}(t)||^2}{\hat{\kappa}(t)^2||\hat{p}(t)||}\right).$$

We then would like to explore the term $||\hat{p}(t)-\hat{\xi}(t)||^2_{G_{pp}(\hat{y}(t),\hat{p}(t))}$ by taking advantage of G_{pp}'s expression.

$$
\begin{aligned}
||\hat{p}(t)-\hat{\xi}(t)||^2_{G_{pp}(\hat{y}(t),\hat{p}(t))} &= (\hat{p}(t)-\hat{\xi}(t))^T\left(G_{pp}(\hat{y}(t),\hat{p}(t))\right)(\hat{p}(t)-\hat{\xi}(t)) \\
&= \frac{v(\hat{y}(t))}{||\hat{p}(t)||^3}(\hat{p}(t)-\hat{\xi}(t))^T\left(||\hat{p}(t)||^2 I-\hat{p}(t)\hat{p}(t)^T\right)(\hat{p}(t)-\hat{\xi}(t)) \\
&= \frac{v(\hat{y}(t))}{||\hat{p}(t)||}(-\hat{\xi}(t)+\hat{\kappa}(t)\hat{p}(t))^T(\hat{p}(t)-\hat{\xi}(t)) \\
&= \frac{v(\hat{y}(t))}{||\hat{p}(t)||}(-\hat{\Xi}(t))^T\left((1-\hat{\kappa}(t))\hat{p}(t)-\hat{\Xi}(t)\right) \\
&= v(\hat{y}(t))\frac{||\hat{\Xi}(t)||^2}{||\hat{p}(t)||}. \qquad \text{(A.77)}
\end{aligned}
$$

Then,

$$
\frac{1}{2}||\hat{p}(t)-\hat{\xi}(t)||^2_{\hat{G}_{pp}} = \frac{v(\hat{y}(t))}{2}\frac{||\hat{\Xi}(t)||^2}{||\hat{p}(t)||}. \qquad \text{(A.78)}
$$

Compare $\hat{\phi}_1(t)$ with equation (A.77),

$$
\begin{aligned}
&v(\hat{x}(t))\left(\frac{1}{2}\frac{||\hat{\Xi}(t)||^2}{\hat{\kappa}(t)^2||\hat{p}(t)||}\right)-\frac{1}{2}v(\hat{y}(t))\frac{||\hat{\Xi}(t)||^2}{||\hat{p}(t)||} \approx -\frac{1}{2}\left(v(\hat{y}(t))-\frac{v(\hat{x}(t))}{\hat{\kappa}(t)^2}\right)\frac{||\hat{\Xi}(t)||^2}{||\hat{p}(t)||} \\
&\quad = O(||\hat{y}(t)-\hat{x}(t)||). \qquad \text{(A.79)}
\end{aligned}
$$

This is because $\hat{\kappa}(t)\sim 1+O(||\hat{p}(t)||^{-\frac{1}{2}})$ and $\hat{\Xi}(t)\sim O(\sqrt{||\hat{p}(t)||})$ followed by $||\hat{p}(t)-\hat{\xi}(t)||\sim O(||p||^{1/2})$ in Lemma 3.3.2.

Then the derivative $\hat{\phi}_1$ becomes,

$$
\hat{\phi}_1(t;x,\xi,y,p)\approx\frac{1}{2}||\hat{p}(t)-\hat{\xi}(t)||^2_{G_{pp}}
$$

$\square$

Proof of Proposition 3.3.10

Proposition A.0.6. *There exists a linear map $\hat{\mathcal{J}}(\hat{t}_c; y, p)$, such that*

$$\begin{bmatrix} \hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c) \\ \hat{p}(\hat{t}_c) - \hat{\xi}(\hat{t}_c) \end{bmatrix} \approx \hat{\mathcal{J}}(\hat{t}_c; y, p) \begin{bmatrix} y - x \\ p - \xi \end{bmatrix}. \tag{A.80}$$

Proof. We consider $(\hat{y}(t) - \hat{x}(t), \hat{p}(t) - \hat{\xi}(t))$ as a function about t and its initial condition $(y - x, p - \xi)$. We denote this initial condition as $(\Delta x, \Delta \xi)$ and $\hat{\mathcal{J}}(t; y, p)$ as the dynamical variational system,

$$\begin{bmatrix} \hat{y}(\hat{t}_c) - \hat{x}(\hat{t}_c; \Delta x, \Delta \xi) \\ \hat{p}(\hat{t}_c) - \hat{\xi}(\hat{t}_c; \Delta x, \Delta \xi) \end{bmatrix} \approx \hat{\mathcal{J}}(\hat{t}_c; y, p) \begin{bmatrix} \Delta x \\ \Delta \xi \end{bmatrix}. \tag{A.81}$$

It is well known that $\hat{\mathcal{J}}(t; y, p)$ satisfies,

$$\frac{d\hat{\mathcal{J}}(t)}{dt} = \begin{bmatrix} G_{xp}(\hat{y}(t), \hat{p}(t)) & G_{pp}(\hat{y}(t), \hat{p}(t)) \\ -G_{xx}(\hat{y}(t), \hat{p}(t)) & -G_{xp}^T(\hat{y}(t), \hat{p}(t)) \end{bmatrix} \hat{\mathcal{J}}(t), \tag{A.82}$$

and its initial condition is an identity matrix,

$$\hat{\mathcal{J}}(0) = I_{2d}. \tag{A.83}$$

By solving $\hat{\mathcal{J}}(t)$, we can get

$$\begin{bmatrix} \hat{y}(t) - \hat{x}(t; y-x, p-\xi) \\ \hat{p}(t) - \hat{\xi}(t; y-x, p-\xi) \end{bmatrix} = \hat{\mathcal{J}}(t; y, p) \begin{bmatrix} y-x \\ p-\xi \end{bmatrix}. \tag{A.84}$$

The transform $\hat{\mathcal{J}}$ defined above is invertible due to the uniqueness of ODE system's solution. □

Proof of Lemma 3.3.9

Lemma A.0.4. *The result after taking the integral about ω can be approximated,*

$$-\int e^{i(\omega-\hat{\tau}_t(\hat{t}_c;y,p))\Delta\hat{t}_0(x,\xi;y,p)} e^{i(\omega-\tilde{\tau}_t(\tilde{t}_c;y,q))\Delta\tilde{t}_0(x,\eta;y,q)} \omega^2 \hat{B}(x,\xi,\omega;y,p)\tilde{B}(x,\eta,\omega;y,q)d\omega$$
$$= e^{iO(1)} K(p,q,y)\hat{\mathcal{B}}(x,\xi;y,p,q)\tilde{\mathcal{B}}(x,\eta;y,p,q) e^{-\frac{i}{2}||\hat{y}(\hat{t}_c)-\hat{x}(\hat{t}_c)||^2_{Re(\hat{M}(\hat{t}_c))}}$$
$$e^{-\frac{i}{2}||\tilde{y}(\tilde{t}_c)-\tilde{x}(\tilde{t}_c)||^2_{Re(\tilde{M}(\tilde{t}_c))}}.$$

Proof. All terms containing ω in $\hat{B}(x,\xi,\omega;y,p)$ (3.110) are as the following,

$$\exp\left(-Im(\hat{\beta})|\omega - \hat{\tau}_t(\hat{t}_c;y,p)|^2 - Im(\hat{\gamma})|\omega - \hat{\tau}_t(\hat{t}_c;x,\xi)|^2\right) \approx$$
$$\exp\left(-Im(\hat{\beta})\left(2\left(\omega - \frac{\hat{\tau}_t(\hat{t}_c;y,p)+\hat{\tau}_t(\hat{t}_c;x,\xi)}{2}\right)^2 + (\hat{\tau}_t(\hat{t}_c;y,p) - \hat{\tau}_t(\hat{t}_c;x,\xi))^2\right)\right)$$
$$\approx e^{-2Im(\hat{\beta})\left(\omega-\frac{\hat{\tau}_t(\hat{t}_c;y,p)+\hat{\tau}_t(\hat{t}_c;x,\xi)}{2}\right)^2} e^{-Im(\hat{\beta})|\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\tau}_t(\hat{t}_c;x,\xi)|^2}. \tag{A.85}$$

Similarly, for $\tilde{B}$,

$$\exp\left(-Im(\tilde{\beta})|\omega-\tilde{\tau}_t(\tilde{t}_c;y,q)|^2-Im(\tilde{\gamma})|\omega-\tilde{\tau}_t(\tilde{t}_c;x,\eta)|^2\right)$$
$$\approx e^{-2Im(\tilde{\beta})\left(\omega-\frac{\tilde{\tau}_t(\tilde{t}_c;y,q)+\tilde{\tau}_t(\tilde{t}_c;x,\eta)}{2}\right)^2} e^{-Im(\tilde{\beta})|\tilde{\tau}_t(\tilde{t}_c;y,q)-\tilde{\tau}_t(\tilde{t}_c;x,\eta)|^2}. \tag{A.86}$$

Combine the terms containing ω in equations (A.85) and (A.86) with $-\omega^2$, and denote

$$-\aleph(\hat{t}_c,\tilde{t}_c;x,\xi,\eta,y,p,q)=\int d\omega e^{-2Im(\tilde{\beta})\left(\omega-\frac{\tilde{\tau}_t(\tilde{t}_c;y,q)+\tilde{\tau}_t(\tilde{t}_c;x,\eta)}{2}\right)^2}\omega^2$$
$$e^{-2Im(\hat{\beta})\left(\omega-\frac{\hat{\tau}_t(\hat{t}_c;y,p)+\hat{\tau}_t(\hat{t}_c;x,\xi)}{2}\right)^2} e^{i(\omega-\hat{\tau}_t(\hat{t}_c;y,p))\Delta\hat{t}_0(x,\xi;y,p)} e^{i(\omega-\tilde{\tau}_t(\tilde{t}_c;y,q))\Delta\tilde{t}_0(x,\eta;y,q)}. \tag{A.87}$$

Then the target integral becomes,

$$-\int e^{i(\omega-\hat{\tau}_t(\hat{t}_0;y,p))\Delta\hat{t}_0(x,\xi;y,p)} e^{i(\omega-\tilde{\tau}_t(\tilde{t}_c;y,q))\Delta\tilde{t}_0(x,\eta;y,q)}\omega^2\hat{B}(x,\xi,\omega;y,p)\tilde{B}(x,\eta,\omega;y,q)d\omega$$
$$=\aleph e^{-Im(\tilde{\beta})|\tilde{\tau}_t(\tilde{t}_c;y,q)-\tilde{\tau}_t(\tilde{t}_c;x,\eta)|^2} e^{-Im(\hat{\beta})|\hat{\tau}_t(\hat{t}_c;y,p)-\hat{\tau}_t(\hat{t}_c;x,\xi)|^2}, \tag{A.88}$$

without considering the constant terms in $\hat{B}$ and $\tilde{B}$. If we can approximate $\aleph$ by the product of a constant $K(p,q,y)$, functions on the receiver side and functions on the source side, then the proposition is proved.

Compute the expression (A.87),

$$-\aleph(\hat{t}_c,\tilde{t}_c;x,\xi,\eta,y,p,q) = e^{-i\hat{\tau}_t(\hat{t}_c;y,p)\Delta\hat{t}_0(x,\xi;y,p)-i\tilde{\tau}_t(\tilde{t}_c;y,p)\Delta\tilde{t}_0(x,\eta;y,q)}e^{-2\frac{Im(\hat{\beta})Im(\tilde{\beta})S_2^2}{Im(\hat{\beta})+Im(\tilde{\beta})}}$$

$$\int \omega^2 e^{i\omega(\Delta\hat{t}_0(x,\xi;y,p)+\Delta\tilde{t}_0(x,\eta;y,q))}e^{-2(Im(\hat{\beta})+Im(\tilde{\beta}))(\omega-S_1)^2}d\omega, \tag{A.89}$$

where

$$S_1(\hat{t}_c,\tilde{t}_c;x,\xi,\eta,y,p,q) = \frac{Im(\hat{\beta})\left(\frac{\hat{\tau}_t(\hat{t}_c;y,p)+\hat{\tau}_t(\hat{t}_c;x,\xi)}{2}\right)+Im(\tilde{\beta})\left(\frac{\tilde{\tau}_t(\tilde{t}_c;y,q)+\tilde{\tau}_t(\tilde{t}_c;x,\eta)}{2}\right)}{Im(\hat{\beta})+Im(\tilde{\beta})}, \tag{A.90}$$

$$S_2(\hat{t}_c,\tilde{t}_c;x,\xi,\eta,y,p,q) = \frac{\hat{\tau}_t(\hat{t}_c;y,p)+\hat{\tau}_t(\hat{t}_c;x,\xi)}{2} - \frac{\tilde{\tau}_t(\tilde{t}_c;y,q)+\tilde{\tau}_t(\tilde{t}_c;x,\eta)}{2}. \tag{A.91}$$

We then have,

$$\aleph = e^{-i\hat{\tau}_t(\hat{t}_c;y,p)\Delta\hat{t}_0(x,\xi;y,p)-i\tilde{\tau}_t(\tilde{t}_c;y,p)\Delta\tilde{t}_0(x,\eta;y,q)}e^{-2\frac{Im(\hat{\beta})Im(\tilde{\beta})S_2^2}{Im(\hat{\beta})+Im(\tilde{\beta})}}$$

$$\frac{d^2\left(e^{iS_1t}e^{-\frac{t^2}{2Im(\hat{\beta}+\tilde{\beta})}}\right)}{dt^2}\Bigg|_{t=\Delta\hat{t}_0(x,\xi;y,p)+\Delta\tilde{t}_0(x,\eta;y,q)}. \tag{A.92}$$

Proposition A.0.7. *Both S_1 and S_2 can be approximated as constants only related to* *the fixed parameter (y,p) and (y,q), that is*

$$S_1(\hat{t}_c,\tilde{t}_c;x,\xi,\eta,y,p,q) \approx \frac{Im(\hat{\beta})\left(\hat{\tau}_t(\hat{t}_c;y,p)\right) + Im(\tilde{\beta})\left(\tilde{\tau}_t(\tilde{t}_c;y,q)\right)}{Im(\hat{\beta}) + Im(\tilde{\beta})},$$

$$S_2(\hat{t}_c,\tilde{t}_c;x,\xi,\eta,y,p,q) \approx \hat{\tau}_t(\hat{t}_c;y,p) - \tilde{\tau}_t(\tilde{t}_c;y,q).$$

Proof. For S_1,

$$\begin{aligned} S_1(\hat{t}_c,\tilde{t}_c;x,\xi,\eta,y,p,q) &= \frac{Im(\hat{\beta})\left(\hat{\tau}_t(\hat{t}_c;y,p)\right) + Im(\tilde{\beta})\left(\tilde{\tau}_t(\tilde{t}_c;y,q)\right)}{Im(\hat{\beta}) + Im(\tilde{\beta})} \\ &+ \frac{Im(\hat{\beta})(\hat{\tau}_t(\hat{t}_c;x,\xi) - \hat{\tau}_t(\hat{t}_c;y,p)) + Im(\tilde{\beta})(\tilde{\tau}_t(\tilde{t}_c;x,\eta) - \tilde{\tau}_t(\tilde{t}_c;y,q))}{2Im(\hat{\beta}) + 2Im(\tilde{\beta})} \\ &= \frac{Im(\hat{\beta})\left(\hat{\tau}_t(\hat{t}_c;y,p)\right) + Im(\tilde{\beta})\left(\tilde{\tau}_t(\tilde{t}_c;y,q)\right)}{Im(\hat{\beta}) + Im(\tilde{\beta})} + O(\sqrt{||\hat{p}_t||}) + O(\sqrt{||\tilde{q}_t||}) \\ &\approx \frac{Im(\hat{\beta})\left(\hat{\tau}_t(\hat{t}_c;y,p)\right) + Im(\tilde{\beta})\left(\tilde{\tau}_t(\tilde{t}_c;y,q)\right)}{Im(\hat{\beta}) + Im(\tilde{\beta})}, \end{aligned} \tag{A.93}$$

since the term in the last step above is about $O(||p|| + ||q||)$, and

$$\begin{aligned} &\frac{Im(\hat{\beta})(\hat{\tau}_t(\hat{t}_c;x,\xi) - \hat{\tau}_t(\hat{t}_c;y,p)) + Im(\tilde{\beta})(\tilde{\tau}_t(\tilde{t}_c;x,\eta) - \tilde{\tau}_t(\tilde{t}_c;y,q))}{2Im(\hat{\beta}) + 2Im(\tilde{\beta})} \\ &\leq |\hat{\tau}_t(\hat{t}_c;x,\xi) - \hat{\tau}_t(\hat{t}_c;y,p)| + |(\tilde{\tau}_t(\tilde{t}_c;x,\eta) - \tilde{\tau}_t(\tilde{t}_c;y,q)| \leq O(\sqrt{||\hat{p}_t|| + ||\tilde{q}_t||}). \end{aligned} \tag{A.94}$$

$$S_1(\hat{t}_c,\tilde{t}_c;x,\xi,\eta,y,p,q) \approx \frac{Im(\hat{\beta})\left(\hat{\tau}_t(\hat{t}_c;y,p)\right) + Im(\tilde{\beta})\left(\tilde{\tau}_t(\tilde{t}_c;y,q)\right)}{Im(\hat{\beta}) + Im(\tilde{\beta})}. \tag{A.95}$$

For S_2,

$$S_2(\hat{t}_c,\tilde{t}_c;x,\xi,\eta,y,p,q) = \hat{\tau}_t(\hat{t}_c;y,p) - \tilde{\tau}_t(\tilde{t}_c;y,q) + O(\sqrt{||p||}), \tag{A.96}$$

due to the fact that $|\hat{\tau}_t(\hat{t}_c;y,p) - \hat{\tau}_t(\hat{t}_c;x,\xi)| = v(y)||p|| - v(x)||\xi||$ and $||p-\xi|| \sim$

$O(\sqrt{||p||})$. Then,

$$e^{-2\frac{Im(\hat{\beta})Im(\tilde{\beta})S_2^2}{Im(\hat{\beta})+Im(\tilde{\beta})}} \approx e^{-2\frac{Im(\hat{\beta})Im(\tilde{\beta})(\hat{\tau}_t(\hat{t}_c;y,p)-\tilde{\tau}_t(\tilde{t}_c;y,q))^2}{Im(\hat{\beta})+Im(\tilde{\beta})}} e^{O(1)}, \tag{A.97}$$

since $Im(\hat{\beta}) \sim O(\frac{1}{||p||})$ and $Im(\tilde{\beta}) \sim O(\frac{1}{||q||})$. We then approximate S_2 by

$$S_2(\hat{t}_c, \tilde{t}_c; x, \xi, \eta, y, p, q) \approx \hat{\tau}_t(\hat{t}_c; y, p) - \tilde{\tau}_t(\tilde{t}_c; y, q). \tag{A.98}$$

We also obtain $|\hat{\tau}_t(\hat{t}_c; y, p) - \tilde{\tau}_t(\tilde{t}_c; y, q)|$ is around $O(\max(\sqrt{||p||}, \sqrt{||q||}))$. □

Proposition A.0.8.

$$\aleph \approx -S_1^2 e^{-\frac{|\Delta\hat{t}_0(x,\xi;y,p)|^2}{4Im(\hat{\beta}+\tilde{\beta})}} e^{-\frac{|\Delta\tilde{t}_0(x,\eta;y,q)|^2}{4Im(\hat{\beta}+\tilde{\beta})}} e^{-2\frac{Im(\hat{\beta})Im(\tilde{\beta})(S_2)^2}{Im(\hat{\beta})+Im(\tilde{\beta})}} \tag{A.99}$$

Proof. The second order time derivative in equation (A.92),

$$\left(\left(iS_1 - \frac{\Delta\hat{t}_0(x,\xi;y,p) + \Delta\tilde{t}_0(x,\eta;y,q)}{Im(\hat{\beta}+\tilde{\beta})}\right)^2 - \frac{1}{Im(\hat{\beta}+\tilde{\beta})}\right)$$
$$e^{iS_1(\Delta\hat{t}_0(x,\xi;y,p)+\Delta\tilde{t}_0(x,\eta;y,q))} e^{-\frac{|\Delta\hat{t}_0(x,\xi;y,p)+\Delta\tilde{t}_0(x,\eta;y,q)|^2}{2(Im(\hat{\beta}+\tilde{\beta}))}}.$$

We now conduct some asymptotic analysis about the terms in the above equation,

$$S_1 \sim O(||p|| + ||q||), \quad \frac{1}{Im(\hat{\beta}+\tilde{\beta})} \sim O(||p|| + ||q||), \tag{A.100}$$

since $\hat{\beta} = \frac{1}{2\hat{\tau}_{tt}(\hat{t}_c;y,p)}$ and $\tilde{\beta} = \frac{1}{2\tilde{\tau}_{tt}(\tilde{t}_c;y,q)}$. On the other hand,

$$\frac{\Delta\hat{t}_0(x,\xi;y,p) + \Delta\tilde{t}_0(x,\eta;y,q)}{Im(\hat{\beta}+\tilde{\beta})} \sim O\left(\frac{1}{\sqrt{||p||}} + \frac{1}{\sqrt{||q||}}\right) O(||p||+||q||). \tag{A.101}$$

Moreover, we have

$$\begin{aligned}
-\frac{|\Delta\hat{t}_0(x,\xi;y,p) - \Delta\tilde{t}_0(x,\eta;y,q)|^2}{2(Im(\hat{\beta}+\tilde{\beta}))} &= -\frac{|\Delta\tilde{t}_0(x,\eta;y,q)|^2}{4Im(\hat{\beta}+\tilde{\beta})} - \frac{|\Delta\hat{t}_0(x,\xi;y,p)|^2}{4Im(\hat{\beta}+\tilde{\beta})} \\
&+ \frac{|\Delta\hat{t}_0(x,\xi;y,p) - \Delta\tilde{t}_0(x,\eta;y,q)|^2}{4Im(\hat{\beta}+\tilde{\beta})} \\
&= -\frac{|\Delta\tilde{t}_0(x,\eta;y,q)|^2}{4Im(\hat{\beta}+\tilde{\beta})} - \frac{|\Delta\hat{t}_0(x,\xi;y,p)|^2}{4Im(\hat{\beta}+\tilde{\beta})} + O(1).
\end{aligned}$$

Consequently,

$$\begin{aligned}
\aleph \approx -S_1^2 e^{i(S_1-\hat{\tau}_t)(\Delta\hat{t}_0(x,\xi;y,p))} e^{i(S_1-\tilde{\tau}_t)(\Delta\tilde{t}_0(x,\eta;y,q))} e^{-\frac{|\Delta\hat{t}_0(x,\xi;y,p)|^2}{4Im(\hat{\beta}+\tilde{\beta})}} \\
e^{-\frac{|\Delta\tilde{t}_0(x,\eta;y,q)|^2}{4Im(\hat{\beta}+\tilde{\beta})}} e^{-2\frac{Im(\hat{\beta})Im(\tilde{\beta})S_2^2}{Im(\hat{\beta})+Im(\tilde{\beta})}}.
\end{aligned} \tag{A.102}$$

With respect to two real-valued phase terms in equation (A.102),

$$e^{i(S_1-\hat{\tau}_t)(\Delta\hat{t}_0(x,\xi;y,p))} = e^{i\frac{Im(\tilde{\beta})(\hat{\tau}_t(\hat{t}_c;y,p)-\tilde{\tau}_t(\tilde{t}_c;y,q))}{Im(\hat{\beta}+\tilde{\beta})}\Delta\hat{t}_0(x,\xi;y,p)} \sim e^{iO(1)}, \tag{A.103}$$

and

$$e^{i(S_1-\tilde{\tau}_t)(\Delta\tilde{t}_0(x,\eta;y,q))} \sim e^{iO(1)}. \tag{A.104}$$

□

To summarize,

$$\aleph \approx -S_1^2 e^{-\frac{|\Delta\hat{t}_0(x,\xi;y,p)|^2}{4Im(\hat{\beta}+\tilde{\beta})}} e^{-\frac{|\Delta\tilde{t}_0(x,\eta;y,q)|^2}{4Im(\hat{\beta}+\tilde{\beta})}} e^{-2\frac{Im(\hat{\beta})Im(\tilde{\beta})(\hat{\tau}_t(\hat{t}_c;y,p)-\tilde{\tau}_t(\tilde{t}_c;y,q))^2}{Im(\hat{\beta})+Im(\tilde{\beta})}}$$
$$\approx K(p,q,y) e^{-\frac{|\Delta\hat{t}_0(x,\xi;y,p)|^2}{4Im(\hat{\beta}+\tilde{\beta})}} e^{-\frac{|\Delta\tilde{t}_0(x,\eta;y,q)|^2}{4Im(\hat{\beta}+\tilde{\beta})}}. \quad \text{(A.105)}$$

This is exactly the goal (A.88) we want to achieve. □

Appendix B

Proof in Elastic Wave

Proof of Positive Definite Hessian Matrix

Theorem B.0.1. *The imaginary part of the Hessian for every beam preserves the S.P.D property after reflection, if it is neither grazing ray nor evanescent wave after reflection.*

Proof. In this appendix, we consider the reflection happens on the surface $\{\mathbf{x} = (x, y, z) : x = 0\}$ without the loss of generality. We denote the phase function of both P and S-wave as τ. The same rule can be applied to the velocity c. To simplify the presentation, all Hessian matrices mentioned below are about the imaginary part only, if not specified. We assume the reflection point is $\mathbf{x}_0 = (x_0, y_0, z_0)$, and all terms below are defined at this point, if not specified. The last simplification is that we follow the positive Hamiltonian throughout this proof, and the negative Hamiltonian will be treated similarly.

Transformation between Hessian matrices

We first define a new matrix $\tilde{M}$ at the reflection point

$$\tilde{M} = \begin{pmatrix} \tau_{tt} & \tau_{ty} & \tau_{tz} \\ \tau_{ty} & \tau_{yy} & \tau_{yz} \\ \tau_{tz} & \tau_{yz} & \tau_{zz} \end{pmatrix}. \tag{B.1}$$

Therefore, we can define a transform between the Hessian matrix M and $\tilde{M}$ following the certain eikonal equation.

$$\aleph(\tilde{M}) = \begin{bmatrix} \tau_{xx} & \tau_{xy} & \tau_{xz} \\ \tau_{xy} & \tau_{yy} & \tau_{yz} \\ \tau_{xz} & \tau_{yz} & \tau_{zz} \end{bmatrix} = M. \tag{B.2}$$

The terms involved with the variable x are defined by following the certain eikonal equation and in the way shown in Section 4.5. Moreover, $\aleph_P$ means that the transform follows the P-wave eikonal equation and $\aleph_S$ follows the S-wave eikonal equation.

If we can prove the transform $\aleph$ and its inverse transform $\aleph^{-1}$ preserve the S.P.D. property, then the theorem is proved since $M^{new} = \aleph_P(\tilde{M}) = \aleph_P(\aleph_S^{-1}(M))$ or $M^{new} = \aleph_S(\tilde{M}) = \aleph_S(\aleph_P^{-1}(M))$. To prove this, instead of considering two types of matrix in equation (B.2) directly, we would like to base our proof first on the complete matrix M_c.

$$M_c = \begin{bmatrix} \tau_{tt} & \nabla\tau_t^T \\ \nabla\tau_t & M \end{bmatrix},$$

where $\nabla\tau_t$ is the gradient of the phase function's time derivative τ_t and M is the original Hessian. Both $\tilde{M}$ and M are submatrix of M_c.

Proof by Contradiction: $\nabla c = 0$

We start with the simpler case, i.e. $\nabla c = 0$ at the point $\mathbf{x}_0$.

Lemma B.0.5. *If ∇c vanishes at $\boldsymbol{x}_0$, then M_c is a positive semi-definite matrix and with rank-three.*

Proof. We write the complete matrix M_c first,

$$\begin{bmatrix} \tau_{tt} & \nabla\tau_t^T \\ \nabla\tau_t & M \end{bmatrix} = \begin{bmatrix} c^2 \frac{\nabla\tau^T}{|\nabla\tau|} M \frac{\nabla\tau}{|\nabla\tau|} & c\frac{\nabla\tau^T M}{|\nabla\tau|} \\ c\frac{M\nabla\tau}{|\nabla\tau|} & M \end{bmatrix}. \tag{B.3}$$

To show that the matrix M_c is positive semi-definite, we use $\boldsymbol{v} = (\alpha, \boldsymbol{p})^T$,

$$\mathbf{v}^T M_c \mathbf{v} = \alpha^2 c^2 \frac{\nabla\tau^T}{|\nabla\tau|} M \frac{\nabla\tau}{|\nabla\tau|} + 2\alpha c \frac{\nabla\tau^T M \boldsymbol{p}}{|\nabla\tau|} + \mathbf{p}^T M \mathbf{p}, \tag{B.4}$$

$$= \left(\alpha c \frac{\nabla\tau}{|\nabla\tau|} + \mathbf{p}\right)^T M \left(\alpha c \frac{\nabla\tau}{|\nabla\tau|} + \mathbf{p}\right). \tag{B.5}$$

Equation (B.5) shows that the null space of M_c is an one-dimensional space and its basis is $\tilde{\boldsymbol{v}}$,

$$\tilde{\boldsymbol{v}} = \begin{pmatrix} \frac{1}{\sqrt{1+c^2}} \\ -\frac{c}{\sqrt{1+c^2}} \frac{\nabla\tau}{|\nabla\tau|} \end{pmatrix}. \tag{B.6}$$

□

The assumption that there are no grazing rays guarantees that $\tau_x \neq 0$ for both beams before and after reflection.

We now start to prove the transform $\aleph$ and its inverse transform preserve the S.P.D. property when $\nabla c = 0$. In other words, if M is S.P.D, then $\tilde{M} = \aleph^{-1}(M)$ is S.P.D. On the other hand, if we have $\tilde{M}$ is S.P.D, then $\aleph(\tilde{M})$ is also S.P.D. for both P-wave and S-wave eikonal equations.

Case I: $\tilde{M} = \aleph^{-1}(M)$

There are several steps involved to prove $\tilde{M}$ is S.P.D. We first consider $\tilde{M}$ is a submatrix of the corresponding complete matrix M_c. Then we use Lemma B.0.5 to show this submatrix is S.P.D.

For any vector $\boldsymbol{u} = (u_1, u_2, u_3) \in \mathbb{R}^3$,

$$\boldsymbol{u}^T \tilde{M} \boldsymbol{u} = \begin{pmatrix} u_1 & 0 & u_2 & u_3 \end{pmatrix} M_c \begin{pmatrix} u_1 \\ 0 \\ u_2 \\ u_3 \end{pmatrix}, \tag{B.7}$$

if the reflection happens on the surface $\{\boldsymbol{x} = (x, y, z) : x = 0\}$. Since there's only a single basis $\tilde{\mathbf{v}}$ in the null space of the matrix M_c, all vectors in the form $(u_1, 0, u_2, u_3)^T$ is not parallel to $\tilde{\boldsymbol{v}}$. Moreover, the complete matrix M_c is a positive semi-definite matrix. Then for any vector $\{\boldsymbol{u} = (u_1, \cdots, u_4) : u_2 = 0\}$, equation (B.7) will be positive. Consequently, $\boldsymbol{u}^T \tilde{M} \boldsymbol{u} > 0$ for any $\boldsymbol{u}$ and $\tilde{M}$ is S.P.D.

Case 2: $M = \aleph(\tilde{M})$

Similar idea will be applied here. The Hessian M firstly is treated as a submatrix of the complete matrix M_c and then use the fact that M_c is a positive semi-definite matrix.

In terms of the transform $M = \aleph(\tilde{M})$, for any vector $\boldsymbol{u} \in \mathbb{R}^3$,

$$\boldsymbol{u}^T M \boldsymbol{u} = \begin{pmatrix} 0 & \boldsymbol{u}^T \end{pmatrix} M_c \begin{pmatrix} 0 \\ \boldsymbol{u} \end{pmatrix}. \tag{B.8}$$

Obviously, all the vectors concerned above $(0, \boldsymbol{u})^t$ is not in the null space of the complete matrix M_c. In other words, for any vector $(0, \boldsymbol{u})$,

$$(0, \boldsymbol{u})^T \neq \beta \tilde{\boldsymbol{v}}, \quad \forall \beta \in \mathbb{R}_{\neq 0} \tag{B.9}$$

Consequently, $\boldsymbol{u}^T M \boldsymbol{u} > 0$ for any $\boldsymbol{u}$ and M is positive definite. $\square$

Proof by Contradiction: $\nabla c \neq 0$

We will follow the similar path as the constant velocity case. First, we prove the imaginary part of the complete matrix M_c is a positive semi-definite matrix.

Lemma B.0.6. *The imaginary part of the complete matrix M_c is a positive semi-definite matrix and with rank-three. Moreover, the single basis $\tilde{\mathbf{v}}$ in its null space is*

$$\tilde{\boldsymbol{v}} = \begin{pmatrix} 1 \\ -c\frac{\nabla \tau}{|\nabla \tau|} \end{pmatrix} \tag{B.10}$$

Proof. We first prove $\tilde{\boldsymbol{v}}$ is in the null space

$$Im(M_c) = \begin{pmatrix} c^2 \frac{\nabla\tau^T Im(M) \nabla\tau}{|\nabla\tau|^2} & c\frac{\nabla\tau^T Im(M)}{|\nabla\tau|} \\ c\frac{Im(M)\nabla\tau}{|\nabla\tau|} & Im(M) \end{pmatrix} \tag{B.11}$$

Here, although the gradient of the velocity $\nabla c \neq 0$, this will only affect the real part. Therefore, we can apply the same argument in Lemma B.0.5 to prove. □

To prove that the transform $\aleph$ and its inverse transform will preserve S.P.D property, we can use the same idea in the case $\nabla c = 0$. The reason is that we only care about the imaginary part of the matrix and their imaginary parts are exactly the same thing as the ones in constant velocity case.

FDTD

As we mentioned previously, the reference solution is generated by the FDTD solution with the staggered grid. Its correctness will be checked here.

To test this, we compare the FDTD solution with the exact solution in the general boundary value problem. If this more general problem is solved correctly, then our reference solution is justified. The parameters used here are,

$$\lambda = 2;$$

$$\mu = 1;$$

$$\rho = 1;$$

while the initial condition $\boldsymbol{f}$ here is zero vector $\overrightarrow{\mathbf{0}}$, and the initial velocity $\boldsymbol{g}$ is

$$\begin{pmatrix} 0 \\ 8\pi\cos(8\pi t)\sin(8\pi x) \\ 0 \end{pmatrix}$$

With the appropriate boundary condition, we know its exact solution is $\sin(8\pi t)\sin(8\pi x)$. We compare the result at $T = 0.8$, In Figure B.1, the blue line is the correct result,

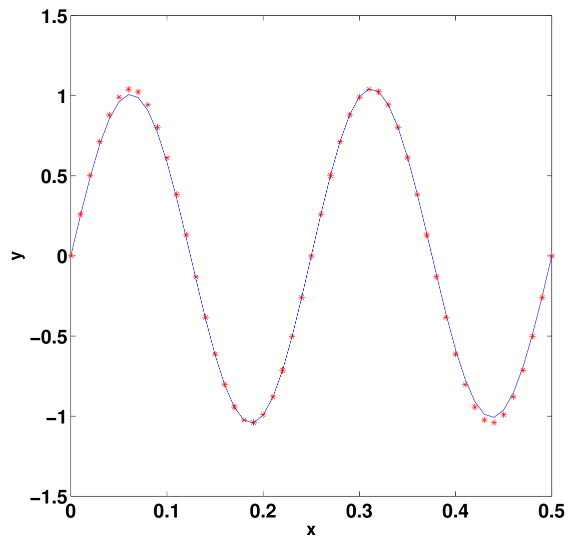


Figure B.1: FDTD solution justify.

while the red star curve is the FDTD result with mesh size $h = 0.01$. Furthermore, we display its convergence rate in Figure B.2.

We start the mesh size from $\frac{1}{50}$ to $\frac{1}{400}$ and each time the grid size is reduced by half. The blue line in Figure B.2 shows the logarithm of L_2-error on each mesh size, while the red star line is a linear function with the slope $\log(1/2)$ for comparison.

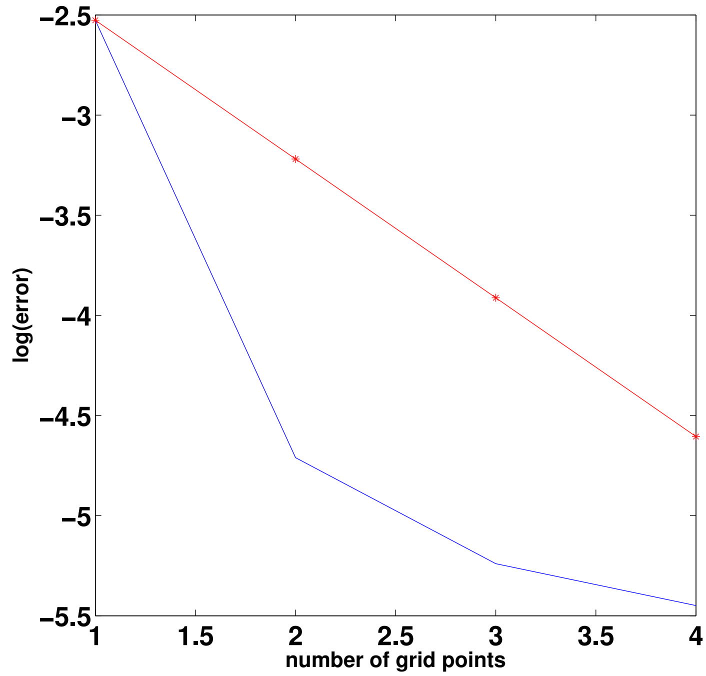


Figure B.2: Convergence Rate of FDTD algorithm

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] K. Aki and P. Richards. *Quantitative seismology.* Freeman and Co., 1980.

[2] U. Albertin, D. Yingst, P. Kitchenside, and V. Tcheverda. True-amplitude beam migration. In *74th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 949–952. Soc. Expl. Geophys., 2004.

[3] V. M. Babich and V. S. Buldyrev. *Asymptotic methods in short wave diffraction problems (in Russian).* Nauka, Moscow, 1972.

[4] G. Bao, J. Lai, and J. Qian. Fast multiscale Gaussian beam methods for wave equations in bounded domains. *J. Comput. Phys.*, 261:36–64, 2014.

[5] G. Bao, J. Qian, L. Ying, and H. Zhang. A convergent multiscale Gaussian-beam parametrix for the wave equation. *Commun. PDEs*, 38:92–134, 2013.

[6] J. D. Benamou. An introduction to Eulerian geometrical optics (1992 - 2002). *J. Sci. Comp.*, 19:63–93, 2003.

[7] G. Beylkin and R. Burridge. Linearized inverse scattering problem of acoustics and elasticity. *Wave Motion*, 12:15–22, 1990.

[8] N. Bleistein. On imaging of reflectors in the earth. *Geophysics*, 52:931–942, 1987.

[9] S. Bougacha, J. Akian, and R. Alexandre. Gaussian beams summation for the wave equation in a convex domain. *Commun. Math. Sci.*, 7:973–1008, 2009.

[10] R. Burridge. Asymptotic evaluation of integrals related to time-dependent fields near caustics. *SIAM J. Appl. Math.*, 55:390–409, 1995.

[11] E. Candès and L. Demanet. The curvelet representation of wave propagators is optimally sparse. *Commun. Pure Appl. Math.*, 58:1472–1528, 2005.

[12] E. Candes, L. Demanet, and L. Ying. Fast computation of Fourier integral operators. *SIAM J. Sci. Comput.*, 29:2464–2493, 2007.

[13] E. Candès and D. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Commun. Pure Appl. Math.*, 57:219–266, 2004.

[14] V. Cerveny. *Seismic Ray Theory.* Cambridge University Press, Cambridge, 2001.

[15] V. Cerveny, M. Popov, and I. Psencik. Computation of wave fields in inhomogeneous media-Gaussian beam approach. *Geophys. J. R. Astr. Soc.*, 70:109–128, 1982.

[16] J. Claerbout. Towards a unified theory of reflection mapping. *Geophysics*, 36:467–481, 1971.

[17] J. F. Claerbout. *Imaging the earth's interior.* Blackwell Scientific Publications, Cambridge, MA, 1985.

[18] M. G. Crandall and P. L. Lions. Viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 277:1–42, 1983.

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

[20] L. Demanet and L. Ying. Wave atoms and time upscaling of wave equations. *Numerische Mathematik*, 113(1):1–71, 2009.

[21] Feng Deng and George A. McMechan. True-amplitude prestack depth migration. *GEOPHYSICS*, 72(3):S155–S166, 2007.

[22] B. Engquist and O. Runborg. Computational high frequency wave propagation. *Acta Numerica*, 12:181–266, 2003.

[23] D. Geba and D. Tataru. A phase space transform adapted to the wave equation. *Commun. Partial Diff. Eq.*, 32:1065–1101, 2007.

[24] Yu Geng, Ru-Shan Wu, and Jinghuai Gao. Gabor-frame-based gaussian packet migration. *Geophysical Prospecting*, 62(6):1432–1452, 2014.

[25] S. Geoltrain and J. Brac. Can we image complex structures with first-arrival traveltime. *Geophysics*, 58:564–575, 1993.

[26] S. Gray. Gaussian beam migration of common shot records. *Geophysics*, 70:S71–S77, 2005.

[27] S. Gray and W. May. Kirchhoff migration using eikonal equation traveltimes. *Geophysics*, 59:810–817, 1994.

[28] N. Hill. Gaussian beam migration. *Geophysics*, 55:1416–1428, 1990.

[29] N. Hill. Prestack Gaussian-beam depth migration. *Geophysics*, 66:1240–1250, 2001.

[30] L. Hörmander. On the existence and the regularity of solutions of linear pseudo-differential equations. *L'Enseignement Mathematique*, 17:99–163, 1971.

[31] S. Leung and J. Qian. Eulerian Gaussian beam methods for Schrödinger equations in the semi-classical regime. *J. Comput. Phys.*, 228:2951–2977, 2009.

[32] S. Leung and J. Qian. The backward phase flow and FBI-transform-based Eulerian Gaussian beams for the Schrödinger equation. *J. Comput. Phys.*, 229:8888–8917, 2010.

[33] S. Leung, J. Qian, and R. Burridge. Eulerian Gaussian beams for high frequency wave propagation. *Geophysics*, 72:SM61–SM76, 2007.

[34] S. Leung, J. Qian, and S. Serna. Fast Huygens Sweeping methods for Schrodinger equations in the semiclassical regime. *Methods Appl. Analy.*, 21:031–066, 2014.

[35] D. Ludwig. Uniform asymptotic expansions at a caustic. *Commun. Pure Appl. Math.*, 19:215–250, 1966.

[36] S. Luo and J. Qian. Factored singularities and high-order Lax-Friedrichs sweeping schemes for point-source traveltimes and amplitudes. *J. Comput. Phys.*, 230:4742–4755, 2011.

[37] S. Luo, J. Qian, and R. Burridge. Fast Huygens sweeping methods for Helmholtz equations in inhomogeneous media. submitted:xxx, 2012.

[38] S. Luo, J. Qian, and R. Burridge. Fast Huygens sweeping methods for Helmholtz equations in inhomogeneous media in the high frequency regime. *J. Comput. Phys.*, 270:378–401, 2014.

[39] S. Luo, J. Qian, and H.-K. Zhao. Higher-order schemes for 3-D traveltimes and amplitudes. *Geophysics*, 77:T47–T56, 2012.

[40] V. P. Maslov. *The Complex WKB Method for Nonlinear Equations I: Linear theory.* Birkhauser Verlag, Basel, 1994.

[41] V. P. Maslov and M. V. Fedoriuk. *Semi-classical approximation in quantum mechanics.* D. Reidel Publishing Company, 1981.

[42] M. Motamed and O. Runborg. Taylor expansion and discretization errors in Gaussian beam superposition. *Wave Motion*, 47:421–439, 2010.

[43] D. Nichols. *Imaging complex structures using band limited Green's functions.* PhD thesis, Stanford University, Stanford, CA94305, 1994.

[44] C. Nolan and W. W. Symes. Global solution of a linearized inverse problem for the wave equation. *Commun. Partial Diff. Eq.*, 22:919–952, 1997.

[45] M. M. Popov. A new method of computation of wave fields using Gaussian beams. *Wave Motion*, 4:85–97, 1982.

[46] M.I. Protasov and Vladimir Cheverda. True amplitude imaging by inverse generalized radon transform based on gaussian beam decomposition of the acoustic green's function. *Geophysical Prospecting*, 59:197 – 209, 03 2011.

[47] J. Qian and L. Ying. Fast Gaussian wavepacket transforms and Gaussian beams for the Schrödinger equation. *J. Comput. Phys.*, 229:7848–7873, 2010.

[48] J. Qian and L. Ying. Fast multiscale Gaussian wavepacket transforms and multiscale Gaussian beams for the wave equation. *SIAM J. Multiscale Modeling and Simulation*, 8:1803–1837, 2010.

[49] J. Ralston. Gaussian beams and the propagation of singularities. Studies in partial differential equations. MAA Studies in Mathematics, 23. Edited by W. Littman. pp.206-248., 1983.

[50] H. Smith. A parametrix construction for wave equations with $C^{1,1}$ coefficients. *Ann. Inst. Fourier Grenoble*, 48:797–835, 1998.

[51] N. Tanushev, B. Engquist, and R. Tsai. Gaussian beam decomposition of high frequency wave fields. *J. Comput. Phys.*, 228:8856–8871, 2009.

[52] N. Tanushev, J. Qian, and J. Ralston. Mountain waves and Gaussian beams. *SIAM J. Multiscale Modeling and Simulation*, 6:688–709, 2007.

[53] A. Toflove and S. C. Huganess. *Computational Electrodynamics: The Finite Difference Time Domain Method, Second Editions.* Artech House, Norwood, MA, 2000.

[54] B. S. White. The stochastic caustic. *SIAM J. Appl. Math.*, 44:127–149, 1984.

[55] B. S. White, A. Norris, A. Bayliss, and R. Burridge. Some remarks on the Gaussian beam summation method. *Geophys. J. R. Astr. Soc.*, 89:579–636, 1987.

[56] Yu Zhang, Guanquan Zhang, and Norman Bleistein. True amplitude wave equation migration arising from true amplitude one-way wave equations. *Inverse Problems*, 19(5):1113, 2003.

[57] Yu Zhang, Guanquan Zhang, and Norman Bleistein. Theory of true-amplitude one-way wave equations and true-amplitude common-shot migration. 70, 07 2005.