COMMUN. MATH. SCI.
Vol. 13, No. 8, pp. 1949–1974

GLOBAL GEOMETRICAL OPTICS APPROXIMATION TO THE HIGH FREQUENCY HELMHOLTZ EQUATION WITH DISCONTINUOUS MEDIA*

CHUNXIONG ZHENG†

Abstract. The global geometrical optics method is a new semi-classical approach for the high frequency linear waves proposed by the author in [33]. In this paper, we rederive it in a more concise way. It is shown that the right candidate of solution ansatz for the high frequency wave equations is the extended WKB function, distinct from the WKB function used in the classical geometrical optics approximation. A new and the main contribution of this paper is an interface analysis for the Helmholtz equation when the incident wave is of extended WKB-type. We derive asymptotic expressions for the reflected and/or transmitted propagating waves in the general case. These expressions are valid even when the incident rays include caustic points.

Key words. High frequency waves, global geometrical optics approximation, caustics, WKB analysis, discontinuous media.

AMS subject classifications. 65M25, 78M35.

1. Introduction

Numerical simulation of high frequency waves is a ubiquitous task in many applied fields such as laser optics [26], underwater acoustics [27], and seismic tomography [22]. The feature of high frequency presents a great numerical challenge to any domain-based direct method, mainly because a large number of grid points are necessarily involved to ensure the basic accuracy requirement. The pollution effect [3] due to numerical dispersion makes this situation even worse. The key ingredient to successfully applying a direct method is to design good preconditioners for the resulting indefinite algebraic system. However, it turns out that most preconditioning techniques, which work efficiently for the discrete elliptic equations, generally perform poorly for the discrete time harmonic wave equations, especially in the high frequency regime [6]. Designing a preconditioner whose performance is independent of both mesh size and frequency parameter has become an active research topic over the last decade. The readers are referred to [6] for a review and to [4, 5, 23, 29] for some successful examples.

There is another class of methods used to compute the high frequency wavefield. These methods resort to some asymptotic ansatz and present approximate but simpler models valid in the high frequency regime. The simplest such kind of method is the geometrical optics method [13]. It is intended to seek a solution of the following WKB-ansatz:

$$u(x) = A^\epsilon(x)\exp[iS(x)/\epsilon],\ A^\epsilon = \sum_{j=0}^{\infty}(-i\epsilon)^j A_j, \tag{1.1}$$

where $1/\epsilon$ relates to the large frequency parameter, S is the real-valued phase function, and the A_j are the amplitude functions. Inserting (1.1) into the governing wave

*Received: June 21, 2013; accepted (in revised form): September 5, 2014. Communicated by Shi Jin.
This work is supported by the National Natural Science Foundation grant No. 10371218, and Tsinghua University Initiative Scientific Research Program.
†Department of Mathematical Sciences, Tsinghua University, Beijing 100084, P.R. China (czheng@math.tsinghua.edu.cn).

equation, say,

$$H(x,-i\epsilon\nabla)u=0,\ x\in\mathbb{R}^n, \tag{1.2}$$

and equating the different powers of ϵ, one derives a Hamilton–Jacobi equation for the phase

$$H(x,\nabla S(x))=0 \tag{1.3}$$

and a sequence of transport equations for the amplitudes A_j which can be solved successively after the phase is uniquely determined. In principle, the Hamilton–Jacobi equation (1.3) can be solved by the method of characteristics. However, this solution is not ensured to be valid in a sufficiently large domain that encloses the region of physical interest. Caustics, around which rays collapse and amplitudes blow up, would generally come into being. In this case, the classical geometrical optics method fails to present a globally valid asymptotic approximation.

Actually, the caustics problem stems from the rigid choice of position representation for the wavefield. When projected to the real space, the Lagrangian manifold induced by the Hamiltonian flow becomes singular at caustics. Equivalently, the real space is no longer a valid representative coordinate Lagrangian plane in the presence of caustics. It is known that for each point of a given Lagrangian manifold, there exists a neighborhood which admits a representative coordinate Lagrangian plane. Based on this fact, Maslov developed the canonical operator method [19]. This method presents an asymptotic approximation of the wavefield with uniform accuracy and thus theoretically overcomes the difficulty of caustics. However, a practical but inevitable issue in Maslov's method is setting up a specific partition of unity which fulfills the requirement of single-valued projection. This is, of course, case-dependent, and to solve it, one needs to understand the structure of the associated Lagrangian manifold. This task is by no means a trivial matter in multiple dimensions. From the computational point of view, a method blind to the Lagrangian manifold is much more preferable.

Along this line, the Gaussian beam approach is a successful class of semi-classical methods which resolve the high frequency wavefield with uniform accuracy. This approach was initiated in 1960's (see [2]). Historically, this approach can be classified into two subcategories, thawed-type (see [8]) and frozen-type (see [9]), depending on whether the beam width is tunable or not during the time evolution. A common point shared by all Gaussian beam approaches is that the central curve of each Gaussian beam is exactly a specific classical ray which obeys the Hamiltonian system. The difference lies in their methodology of function approximation. In the thawed-type approaches, each individual beam is an asymptotic solution of the governing wave equation. This renders the beam summation as a suitable asymptotic solution in a large domain. Comparatively, in the frozen-type approaches, each individual beam is not an asymptotic solution. The accuracy of the beam summation is achieved by a delicate choice of prefactor which balances the contribution of each single beam. The Gaussian beam approach is easy to implement, straightforward to obtain higher accuracy [24, 12], and elegant from the methodological point of view. In recent years, there have been many papers addressing the accuracy analysis of this method [20, 14, 15, 16]. Besides, the Eulerian formulation of Gaussian beam approach has been systematically developed [10, 11, 28].

From the computational point of view, Gaussian beam approaches have shortcomings. For thawed-type methods, the beam width expands as the rays evolve. For a long time or large domain simulation, re-initialization is generally needed to ensure the accuracy of numerical solutions, and so far this treatment can only be performed numerically (see [1, 25, 31]). For frozen-type methods, the approximate solutions are expressed as an oscillatory integral defined on a manifold of much higher dimension. For example, the dimension of the integral manifold involved in the well-known Herman–Kluk approximation [9] to the Schrödinger equation is $3N$ for a problem with N degrees of freedom. Rousse [21] proposed a new frozen-type semi-classical approximation to the Schrödinger equation. However, the dimension of his integral manifold, though smaller, is still $2N$. Similar things happen in the frozen-type Gaussian beam approach for the wave equation and strictly hyperbolic system studied by Lu and Yang [17, 18].

Recently, the author proposed a new semi-classical approximation, the global geometrical optics method [33], for solving general scalar linear equations in the high frequency regime. This method can be viewed as an improvement of the classical geometrical optics method. It is blind to the structure of underlying Lagrangian manifold and presents a semi-classical approximation of the wavefield with uniform asymptotic accuracy. Analogous to the Gaussian beam approaches, the wavefield given by the global geometrical optics method is expressed as an integral of Gaussian functions parameterized by some manifold. However, this new semi-classical method owns the merits of both thawed-type and frozen-type methods. On the one hand, the width of Gaussian functions is fixed in frozen-type methods; thus there is no difficulty from beam spreading. On the other hand, the dimension of the integral manifold is minimal in the thawed-type method: it is N for problems with N degrees of freedom.

In this paper, we rederive the global geometrical optics approximation for general linear wave equations in a manner analogous to, but more concise than, our derivation in [33]. We introduce a new concept—the extended WKB function which includes the classical WKB function as a subset. We show that the right candidate of solution ansatz for the high frequency wave equations is the extended WKB function instead of the classical WKB function which results in the difficulty of caustics. Another new and the main contribution of this paper is an interface analysis for the Helmholtz equation when the incident wave is of extended WKB-type. This research is fundamental when extending and applying the global geometrical optics approximation to more complicated piecewise smooth media problems such as seismic migration in computational geophysics. In the case when the wave is totally reflected, or partially reflected and partially transmitted, we show that the split waves are also of extended WKB-type. A subtle issue in this interface analysis is determining geometric factors related to the geometries of the incident and split manifolds when the incident rays include some caustic points. These geometric factors are needed to determine the amplitudes of split waves.

The rest of this paper is organized as follows. In Section 2, we introduce basic notation and collect some results from [33] which are necessary for the later derivation. In Section 3, we rederive the global geometrical optics approximation for the scalar wave equations. In Section 4, we perform an asymptotic interface analysis for the Helmholtz equation, and in Section 5, we discuss some miscellaneous issues related to the global geometrical optics approximation. Finally, in Section 6, we conclude this paper.

2. Preliminaries

We start this section with a brief explanation of notation and definitions.

- We use I_N to denote the identity matrix of dimension N. The canonical symplectic matrix is denoted by

$$J_{2N} = \begin{bmatrix} 0 & I_N \\ -I_N & 0 \end{bmatrix}.$$

The dimension parameter will be omitted whenever it is clear from context.

- Any plane can be represented by a matrix whose columns vectors constitute a basis of the plane. We do not distinguish a plane from its specific matrix representation.
- The symplectic inner product in $\mathbb{R}^{2N}$ is defined as

$$[z,z'] = (z,Jz'),\ \forall z,z' \in \mathbb{R}^{2N}.$$

  A plane is Lagrangian if any two column vectors of its representative matrix are symplectic orthogonal.
- Throughout this paper, unless explicitly specified, we use the subscripts $_1$ and $_2$ to indicate the upper and lower parts of an N-plane in the $2N$-dimensional phase space. This implies that given any N-plane $C \in \mathbb{R}^{2N\times N}$, we have

$$C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix},\ C_1, C_2 \in \mathbb{R}^{N\times N}.$$

  If $\det C_1 \neq 0$ and $C_2 C_1^{-1}$ is similar to a real diagonal matrix, the following definition

$$\mathcal{M}(C) = \det \sqrt{I_N - iC_2 C_1^{-1}} \tag{2.1}$$

  is appropriate since there is no ambiguity in the branch choice of the square root function. In particular, if C is Lagrangian and C_1 is invertible, then $C_2 C_1^{-1}$ is symmetric, and thus $\mathcal{M}(C)$ is well-defined.
- Any N-dimensional Lagrangian plane in $\mathbb{R}^{2N}$ admits a unitary matrix representation. Suppose $C \in \mathbb{R}^{2N\times N}$ is a matrix representation. Let $C = QP$ be the polar decomposition and set $U = Q_1 + iQ_2$. Then U is unitary.

2.1. Heisenberg group and Weyl quantization. The Heisenberg Lie group $\mathbf{H}_N$ is $\mathbb{R}^{2N+1}$ equipped with the group law

$$(z,s)(z',s') = (z+z', s+s'+[z,z']/2),\ \forall z,z' \in \mathbb{R}^{2N},\ \forall s,s' \in \mathbb{R}.$$

The unitary representation of $\mathbf{H}_N$ is defined by

$$\rho_\epsilon(z,s) = \exp\{i([z,W]-s)/\epsilon\}, \quad W = (x, -i\epsilon\nabla).$$

Throughtout this paper, we make the convention that $\rho_\epsilon(z) = \rho_\epsilon(z,0)$. For any $f \in L^2(\mathbb{R}^N)$, it is straightforward to verify that (or consult Page 21 in [7])

$$[\rho_\epsilon(z)f](x) = \exp[-ip(x+q/2)/\epsilon]\, f(x+q). \tag{2.2}$$

For any $H = H(\omega) = H(x,\xi) \in \mathcal{S}'(\mathbb{R}^N_x \times \mathbb{R}^N_\xi)$ with $\omega = (x,\xi)$, the Weyl quantization of H, denoted by $H(W)$, is a linear operator determined by

$$H(W) = H(x, -i\epsilon\nabla) = (2\pi)^{-2N} \int \hat{H}(z)\rho_\epsilon(\epsilon z)dz,$$

where $\hat{H}$ indicates the symplectic Fourier transform of H defined by

$$\hat{H}(z) = \int H(\omega)\exp(i[\omega,z])\,d\omega.$$

2.2. Unitary representation of unitary matrix group. The Bargmann transform $\mathcal{B}_\epsilon$ is defined by

$$[\mathcal{B}_\epsilon f](\varsigma) = \int \exp\left[-(2x^2 - 4\varsigma x + \varsigma^2)/4\epsilon\right] f(x)dx, \ \forall \varsigma \in \mathbb{C}^N.$$

The transform $\mathcal{B}_\epsilon$ is an isometry from $L^2(\mathbb{R}^N)$ into the Fock space

$$\mathcal{F}_N = \left\{ F \text{ entire } \Big| \|F\|^2 = 2^{N/2}(2\pi\epsilon)^{-3N/2} \int |F(\varsigma)|^2 \exp\left[-|\varsigma|^2/2\epsilon\right] d\varsigma < \infty \right\}.$$

The inverse Bargmann transform $\mathcal{B}_\epsilon^{-1}$ from $\mathcal{F}_N$ to $L^2(\mathbb{R}^N)$ is given by

$$[\mathcal{B}_\epsilon^{-1} F](x) = 2^{N/2}(2\pi\epsilon)^{-3N/2} \int \exp\left[-(2x^2 - 4\bar{\varsigma} x + \bar{\varsigma}^2)/4\epsilon\right] \exp\left[-|\varsigma|^2/2\epsilon\right] F(\varsigma) d\varsigma.$$

Let $\mathcal{U}_N$ denote the unitary matrix group acting on $\mathbb{C}^N$. For any $U \in \mathcal{U}_N$, we define

$$[\mathcal{T}_U F](\varsigma) = F(U^\dagger \varsigma), \ \forall F \in \mathcal{F}_N,$$

and

$$\mu_\epsilon(U) = \mathcal{B}_\epsilon^{-1} \mathcal{T}_U \mathcal{B}_\epsilon.$$

Here and hereafter, the symbol † indicates the real transpose operator. Since $\mathcal{T}$ is a unitary representation of $\mathcal{U}_N$ on the Fock space $\mathcal{F}_N$, μ_ϵ is a unitary representation of $\mathcal{U}_N$ on $L^2(\mathbb{R}^N)$.

In the following, we list some properties of the representation operators ρ_ϵ and μ_ϵ. Their proofs can be found in [33] with some slight formulation differences.

LEMMA 2.1. *Define the fundamental coherent state function as*

$$\phi_\epsilon(x) = (2\pi\epsilon)^{-N/2} \exp\left(-x^2/2\epsilon\right).$$

For any unitary matrix $U \in \mathcal{U}_N$, it holds that $\mu_\epsilon(U)\phi_\epsilon = \phi_\epsilon$.

LEMMA 2.2. *For any $U = U_R + iU_I \in \mathcal{U}_N$, $z \in \mathbb{R}^{2N}$, and $H \in \mathcal{S}'(\mathbb{R}_x^N \times \mathbb{R}_\xi^N)$, it holds that*

$$\rho_\epsilon(z) H(W) \rho_\epsilon(-z) = H(W + z), \tag{2.3}$$

$$\mu_\epsilon(U^*) H(W) \mu_\epsilon(U) = H(R_U W), \tag{2.4}$$

where R_U is an orthogonal symplectic matrix defined by

$$R_U = \begin{bmatrix} U_R & -U_I \\ U_I & U_R \end{bmatrix}.$$

LEMMA 2.3. *Suppose $z = z(t)$ is a smooth curve in $\mathbb{R}^{2N}$ and $U = U(t)$ is a smooth curve in $\mathcal{U}_N$. Then it holds that*

$$-i\epsilon\partial_t \rho_\epsilon(z) = ([\dot{z}, W] - [z, \dot{z}]/2)\, \rho_\epsilon(z),$$
$$-i\epsilon\partial_t \mu_\epsilon(U^*) = (W\Phi W - \epsilon \mathrm{tr} T_I)\, \mu_\epsilon(U^*)/2,$$
$$\mu_\epsilon(U^*)\rho_\epsilon(z)(-i\epsilon\partial_t)\left[\rho_\epsilon(-z)\mu_\epsilon(U)\right] = -[\dot{z}, R_U W] - W\Phi W/2 + [z, \dot{z}]/2 + \epsilon \mathrm{tr} T_I/2.$$

In the above equations, T_I and Φ are given by

$$T_R + iT_I = \dot{U}^* U, \ \Phi = \begin{bmatrix} T_I & T_R \\ -T_R & T_I \end{bmatrix}.$$

2.3. Extended WKB function.

DEFINITION 2.4. *A (standard) WKB function is an asymptotic series of the form*

$$u(x) = [A(x) + (-i\epsilon)A_1(x) + \cdots]\exp[iS(x)/\epsilon], \ x \in \mathbb{R}^N.$$

The Lagrangian manifold associated with a WKB function is the graph of the differential dS in the phase space $\mathbb{R}^{2N}$. We use the same notation dS to indicate this Lagrangian manifold; i.e.,

$$dS = \{(q,p) | p = \nabla S(q)\}.$$

A WKB function is said to admit the vanishing derivatives property if

$$\nabla S(0) = 0, \quad \nabla^2 S(0) = 0.$$

The geometric interpretation of the vanishing derivatives property is that the Lagrangian manifold of a WKB function is tangent to the real plane at the origin.

The following formula plays a fundamental role in the classical WKB analysis:

$$H(W)\{A(x)\exp[iS(x)/\epsilon]\} = \exp[iS(x)/\epsilon]\sum_{j=0}^{\infty}(-i\epsilon)^j R_j[A]. \tag{2.5}$$

Here $H(W)$ is the Weyl quantization of the symbol $H = H(w) = H(x,\xi)$. The operators R_j are linear differential operators of order j, with the first two given by

$$R_0[A] = HA, \quad R_1[A] = \nabla_\xi H \cdot \nabla A + \operatorname{tr}\left(\nabla_\xi^2 H \nabla^2 S + \nabla_x \nabla_\xi H\right) A/2.$$

In the above equations, the function H and its derivatives are evaluated at the phase point $(x, \nabla S(x))$.

DEFINITION 2.5. *An extended WKB function is an asymptotic series of the form*

$$u(x) = \int_{z\in\Lambda} [\mathcal{A}(z) + (-i\epsilon)\mathcal{A}_1(z) + \cdots]\exp[i\mathcal{S}(z)/\epsilon][\rho_\epsilon(-z)\phi_\epsilon](x)d\mathrm{vol}, \ x \in \mathbb{R}^N,$$

where Λ is a Lagrangian manifold in the phase space $\mathbb{R}^{2N}$. The amplitude $\mathcal{A}$ (and $\mathcal{A}_i$) and the (real) phase $\mathcal{S}$ are only defined on Λ. Additionally, $\mathcal{S}$ is assumed to be a generating function of the differential 1-form $pdq - d(pq)/2$ on Λ.

Throughout this paper, we are merely concerned with the first order approximation of the relevant quantities. In most cases, only the leading order term of the amplitude series is important. For brevity of notation, we simply use $\mathcal{O}(\epsilon)$ to denote the higher-order terms. The following theorem reveals that the extended WKB function is indeed an extended concept of the standard WKB function. Its proof combines Theorem 3.2 and Lemma 3.3 in [33], and we omit it here.

THEOREM 2.6. *Given an extended WKB function*

$$u(x) = \int_{z\in\Lambda} [\mathcal{A}(z) + \mathcal{O}(\epsilon)]\exp[i\mathcal{S}(z)/\epsilon][\rho_\epsilon(-z)\phi_\epsilon](x)d\mathrm{vol}.$$

For any $z = (q,p) \in \operatorname{supp} A$, let U be a unitary matrix representation of the tangent plane. Then $\mu_\epsilon(U^*)\rho_\epsilon(z)u$ is a local WKB function with the vanishing derivatives property, and it holds that

$$[\mu_\epsilon(U^*)\rho_\epsilon(z)u](0) = [\mathcal{A}(z) + \mathcal{O}(\epsilon)]\exp[i\mathcal{S}(z)/\epsilon].$$

Furthermore, if Λ *admits a local one-to-one projection onto the real plane, then we have*

$$\varphi(q) = [\mathcal{A}(z)\mathcal{M}(C(z)) + \mathcal{O}(\epsilon)]\exp\{i[\mathcal{S}(z) + pq/2]/\epsilon\}.$$

Here $C(z) \in \mathbb{R}^{2N\times N}$ *is any matrix representation of the tangent plane of* Λ *at the point* z.

3. Global geometrical optics approximation

In this section, we derive the global geometrical optics approximation to the high frequency scalar wave equation in its general form

$$H(W)u = 0, \ W = (x, -i\epsilon\nabla). \tag{3.1}$$

Here $H(W)$ denotes the Weyl quantization of the symbol $H(w) = H(x,\xi)$ with $w = (x,\xi)$. The zero level set of H is denoted by $\mathcal{N}$. Suppose $\Lambda_0 \subset \mathcal{N}$ is an isotropic manifold of dimension $N-1$ in the phase space $\mathbb{R}^{2N}$. If the Hamilton vector field $J\nabla H(z)$ is transversal to Λ_0, then the Hamiltonian system

$$\dot{z} = J\nabla H(z), \tag{3.2}$$

equipped with the initial data

$$z|_{t=0} = z_0 \in \Lambda_0,$$

renders a Lagrangian manifold embedded in $\mathcal{N}$. Let us denote this Lagrangian manifold by Λ. Suppose the Hamiltonian H is specified in a way such that Λ is topologically isomorphic to $\Lambda_0 \times \mathbb{R}$. Let y be a coordinate of Λ_0. Then Λ is parameterized by (t,y). Therefore, the Jacobian $C = \partial z/\partial(t,y)$ determined by

$$\dot{C} = J\nabla^2 H(z)C, \quad C|_{t=0} = \left[J\nabla H(z_0)\ \partial_y z_0\right] \tag{3.3}$$

forms a matrix representation of the tangent plane of Λ at the point $z = z(t,y)$. Let

$$C = QP, \ Q \in \mathbb{R}^{2N\times N}, \ P \in \mathbb{R}^{N\times N} \tag{3.4}$$

be the polar decomposition of C; i.e.,

$$P = \left(C^\dagger C\right)^{1/2}, \ Q = C\left(C^\dagger C\right)^{-1/2}.$$

Set

$$U = Q_1 + iQ_2. \tag{3.5}$$

Then U is a unitary matrix.

Given a generating function $\mathcal{S}_0$ of $pdq - d(pq)/2$ on Λ_0, if $\mathcal{S}$ is the solution to the following ODE problem

$$\dot{\mathcal{S}} + [z,\dot{z}]/2 = 0, \quad \mathcal{S}|_{t=0} = \mathcal{S}_0(z_0),$$

then $\mathcal{S}$ is a generating function of $pdq - d(pq)/2$ on Λ. Let us introduce the operator $\mathcal{K}$ as

$$[\mathcal{K}\mathcal{A}](x) = \int_{\varsigma\in\Lambda} \mathcal{A}(\varsigma)\exp[i\mathcal{S}(\varsigma)/\epsilon]\,[\rho_\epsilon(-\varsigma)\phi_\epsilon](x)d\mathrm{vol}.$$

This is an extended WKB function with the Lagrangian manifold Λ. Let (z,U) be any moving frame induced by the underlying Hamiltonian flow and positioned on Λ; i.e., $z=z(t)$ solves the Hamiltonian system (3.2) and $U=U(z(t))$ is determined by (3.3), (3.4), and (3.5). Applying Lemma 2.2 and Lemma 2.3, we derive

$$\begin{aligned}&\mu_\epsilon(U^*)\rho_\epsilon(z)H(W)\mathcal{K}\mathcal{A}=\mu_\epsilon(U^*)\rho_\epsilon(z)[H(W)-i\epsilon\partial_t]\mathcal{K}\mathcal{A}\\&=\mu_\epsilon(U^*)\rho_\epsilon(z)[H(W)-i\epsilon\partial_t]\rho_\epsilon(-z)\mu_\epsilon(U)\mu_\epsilon(U^*)\rho_\epsilon(z)\mathcal{K}\mathcal{A}\\&=\left[\tilde{H}(W,t)+\epsilon\mathrm{tr}T_I/2-i\epsilon\partial_t\right]\mu_\epsilon(U^*)\rho_\epsilon(z)\mathcal{K}\mathcal{A},\end{aligned}$$

where

$$\tilde{H}(w,t)=H(R_Uw+z)-\nabla H(z)\cdot R_Uw-w\Phi w/2+[z,\dot{z}]/2. \tag{3.6}$$

By Theorem 2.6, $\mu_\epsilon(U^*)\rho_\epsilon(z)\mathcal{K}\mathcal{A}$ is a local WKB function with the vanishing derivatives property. Suppose

$$[\mu_\epsilon(U^*)\rho_\epsilon(z)\mathcal{K}\mathcal{A}](x,t)=[A_0(x,t)+\mathcal{O}(\epsilon)]\exp[iS(x,t)/\epsilon]. \tag{3.7}$$

Then we have

$$A_0(0,t)=\mathcal{A}(z), \tag{3.8}$$

and

$$S(0,t)=\mathcal{S}(z),\quad \nabla S(0,t)=0,\quad \nabla^2S(0,t)=0. \tag{3.9}$$

For any smooth function $\varphi=\varphi(x,t)$, applying the fundamental formula (2.5), we derive

$$\left[\tilde{H}(W,t)+\epsilon\mathrm{tr}T_I/2-i\epsilon\partial_t\right][\varphi\exp(iS/\epsilon)]=\exp(iS/\epsilon)\sum_{j=0}^{\infty}(-i\epsilon)^j\mathcal{T}_j[\varphi],$$

where

$$\begin{aligned}&\mathcal{T}_0[\varphi]=\left(\dot{S}+\tilde{H}\right)\varphi,\\&\mathcal{T}_1[\varphi]=\dot{\varphi}+\nabla_\xi\tilde{H}\cdot\nabla\varphi+\left[\nabla_\xi^2\tilde{H}:\nabla^2S+\mathrm{tr}(\nabla_x\nabla_\xi\tilde{H})+i\mathrm{tr}T_I\right]\varphi/2.\end{aligned}$$

In the above equations, the function $\tilde{H}$ and its derivatives are evaluated at the point $(x,\nabla S(x,t),t)$.

Recalling the definition of $\tilde{H}$ (see (3.6)) and the definition of R_U (see Lemma 2.2), we have

$$\tilde{H}(0,t)=[z,\dot{z}]/2,\quad \nabla\tilde{H}(0,t)=0, \tag{3.10}$$

$$\nabla^2\tilde{H}(0,t)=\begin{bmatrix}Q^\dagger\nabla^2H(z)Q & -Q^\dagger\nabla^2H(z)JQ\\ Q^\dagger J\nabla^2H(z)Q & -Q^\dagger J\nabla^2H(z)JQ\end{bmatrix}-\begin{bmatrix}T_I & T_R\\ -T_R & T_I\end{bmatrix}. \tag{3.11}$$

By (3.9) and (3.10), we have

$$\mathcal{T}_0[\varphi]|_{x=0}=\left(\dot{\mathcal{S}}+[z,\dot{z}]/2\right)\varphi(0,t)=0.$$

Considering T_R is skew-symmetric and $T_I = Q^\dagger \nabla^2 H(z) Q$, recalling (3.9), (3.10), and (3.11), we have

$$\mathcal{T}_1[\varphi]|_{x=0} = \dot{\varphi}(0,t) + \operatorname{tr}\left[Q^\dagger (iI+J)\nabla^2 H(z)Q\right]\varphi(0,t)/2 \stackrel{def}{=} \mathcal{L}(\varphi(0,t)).$$

Therefore, recalling (3.7) and (3.8), we derive

$$[\mu_\epsilon(U^*)\rho_\epsilon(z)H(W)\mathcal{K}\mathcal{A}]|_{x=0} = \left[(-i\epsilon)\mathcal{L}(\mathcal{A}) + \mathcal{O}(\epsilon^2)\right]\exp(i\mathcal{S}/\epsilon).$$

Applying Theorem 2.6, we arrive at

$$H(W)\mathcal{K}\mathcal{A} = (-i\epsilon)\mathcal{K}[\mathcal{L}(\mathcal{A}) + \mathcal{O}(\epsilon)]. \tag{3.12}$$

The formula (3.12) implies that if $\mathcal{A}$ is such that $\mathcal{L}(\mathcal{A}) = 0$, i.e.

$$\dot{\mathcal{A}} + \operatorname{tr}\left[Q^\dagger (iI+J)\nabla^2 H(z)Q\right]\mathcal{A}/2 = 0, \tag{3.13}$$

then we have

$$H(W)\mathcal{K}\mathcal{A} = \mathcal{O}(\epsilon^2).$$

For a short conclusion, given an isotropic manifold Λ_0 transversal to the Hamilton vector field $J\nabla H(z)$ and embedded in the null manifold of the Hamiltonian H, a generating function $\mathcal{S}_0$ of $pdq - d(pq)/2$ on Λ_0, and a smooth function $\mathcal{A}_0$ on Λ_0, if we determine a Lagrangian manifold Λ by

$$\dot{z} = J\nabla H(z), \quad z|_{t=0} = z_0 \in \Lambda_0,$$

a generating function $\mathcal{S}$ of $pdq - d(pq)/2$ on Λ by

$$\dot{\mathcal{S}} + [z,\dot{z}]/2 = 0, \quad \mathcal{S}|_{t=0} = \mathcal{S}_0(z_0),$$

and an amplitude function $\mathcal{A}$ on Λ by

$$\begin{aligned} &\dot{C} = J\nabla^2 H(z)C,\ C = QP, \quad C|_{t=0} = \left[J\nabla H(z_0)\ \partial_y z_0\right], \\ &\dot{\mathcal{A}} + \operatorname{tr}\left[Q^\dagger (iI+J)\nabla^2 H(z)Q\right]\mathcal{A}/2 = 0, \quad \mathcal{A}|_{t=0} = \mathcal{A}_0(z_0), \end{aligned}$$

then the extended WKB function

$$u(x) = \int_{z\in\Lambda} \mathcal{A}(z)\exp[i\mathcal{S}(z)/\epsilon][\rho_\epsilon(-z)\phi_\epsilon](x)d\mathrm{vol} \tag{3.14}$$

gives an asymptotic solution of first order to the wave equation (3.1). We call this approximation the global geometrical optics approximation since it is globally valid and analogous to the classical geometrical optics method. The main tool we used is merely WKB analysis, although we performed it with a moving frame technique in the phase space.

4. Helmholtz equation with discontinuous media

We consider the Helmholtz equation

$$-\epsilon^2 \Delta u - s^2 u = 0, \tag{4.1}$$

where $\epsilon = \omega^{-1}$ is the inverse of angular frequency and s is the slowness field, the inverse of the velocity field. The Helmholtz equation (4.1) is a special instance of the general wave equation (3.1) with

$$H(x,\xi) = [\xi^2 - s^2(x)]/2.$$

Suppose standard WKB Cauchy data is specified on a smooth manifold Γ of codimension 1 in the spatial domain

$$u_\Gamma(x) = [A_\Gamma(x) + \mathcal{O}(\epsilon)]\exp[iS_\Gamma(x)/\epsilon],\ \forall x \in \Gamma. \tag{4.2}$$

By a standard WKB analysis, we know that if

$$|\nabla_\Gamma S_\Gamma(x)| \neq s(x),\ \forall x \in \Gamma,$$

there are two local WKB solutions $u_\pm$ around Γ which satisfy (within $\mathcal{O}(\epsilon)$)

$$u_\pm(x) = u_\Gamma(x),\ \forall x \in \Gamma, \tag{4.3}$$
$$(-i\epsilon)\nabla u_\pm(x) = \left[\nabla_\Gamma S_\Gamma(x) \pm \zeta(x)n(x)\right] u_\Gamma(x),\ \forall x \in \Gamma. \tag{4.4}$$

In the above equation, n is the normal direction of Γ and $\zeta(x)$ is the impedance determined by

$$\zeta(x) = \left[s^2(x) - |\nabla_\Gamma S_\Gamma(x)|^2\right]^{1/2},\ \forall x \in \Gamma.$$

Note that if $s(x) < |\nabla_\Gamma S_\Gamma(x)|$, the impedance $\zeta(x)$ is purely imaginary with a positive imaginary part.

Now suppose the slowness field s is piecewise smooth with an interface Γ dividing the spatial domain into two parts, Ω_1 and Ω_2. Let us denote the slowness fields on Ω_1 and Ω_2 by s_1 and s_2, respectively. At the interface Γ, the wavefield is assumed to satisfy the connection conditions

$$u|_{\Gamma,\Omega_2} - u|_{\Gamma,\Omega_1} = 0,\ n^\dagger \nabla u|_{\Gamma,\Omega_2} - n^\dagger \nabla u|_{\Gamma,\Omega_1} = 0. \tag{4.5}$$

We consider an incident wave of extended WKB-type which propagates in Ω_1 and is initiated from an isotropic manifold Λ_0 (parameterized by y) of dimension $N-1$ embedded in the null manifold of the Hamiltonian $H(x,\xi)$. That is to say, we suppose the incident wavefield is given as

$$u_I(x) = \int_{z_I = (q_I, p_I) \in \Lambda_I} \mathcal{A}_I(z_I)\exp[i\mathcal{S}_I(z_I)/\epsilon][\rho_\epsilon(-z_I)\phi_\epsilon](x)d\mathrm{vol}, \tag{4.6}$$

where the incident manifold Λ_I and other relevant quantities are determined by

$$\dot{z}_I = J\nabla H(z_I),\ z_I|_{t=0} = z_0 \in \Lambda_0, \tag{4.7}$$
$$\dot{\mathcal{S}}_I + [z_I, \dot{z}_I]/2 = 0,\ \mathcal{S}_I|_{t=0} = \mathcal{S}_0(z_0), \tag{4.8}$$
$$\dot{C}_I = J\nabla^2 H(z_I)C_I,\ C_I = Q_I P_I,\ C_I|_{t=0} = \left[J\nabla H(z_0)\ \partial_y z_0\right], \tag{4.9}$$
$$\dot{\mathcal{A}}_I + \mathrm{tr}\left[Q_I^\dagger(iI + J)\nabla^2 H(z_I)Q_I\right]\mathcal{A}_I/2 = 0,\ \mathcal{A}_I|_{t=0} = \mathcal{A}_0(z_0). \tag{4.10}$$

In the above equations, $\mathcal{S}_0$ is a generating function of the differential 1-form $pdq - d(pq)/2$ on Λ_0 and $\mathcal{A}_0$ is a prescribed smooth function on Λ_0. Suppose the interface Γ

is specified as the zero-level set of the function $\Phi=\Phi(q)$. Then the *time* $\tau=\tau(y)$ for a ray indexed by y to impinge Γ satisfies

$$\Phi(q_I(\tau,y))=0.$$

Differentiating the above equation with respect to y yields

$$n^\dagger \partial_t q_I(\tau,y)\partial_y \tau + n^\dagger \partial_y q_I(\tau,y)=0.$$

Therefore, we have

$$\partial_y \tau = -\left[n^\dagger \partial_t q_I(\tau,y)\right]^{-1}\left[n^\dagger \partial_y q_I(\tau,y)\right].$$

4.1. Reflected and transmitted waves. Let us set

$$q_\Gamma = q_I(\tau,y),\ p_{I,\Gamma}=p_I(\tau,y),\ z_{I,\Gamma}=(q_\Gamma,p_{I,\Gamma}),\ C_{I,\Gamma}=C_I(z_{I,\Gamma}).$$

When $z_{I,\Gamma}$ is a regular point of the projection map from the Lagrangian manifold Λ_I to the real plane, the matrix $C_{I,\Gamma,1}$ (the upper half part of $C_{I,\Gamma}$ by our convention) is invertible, and $\mathcal{M}(C_{I,\Gamma})$ is well-defined. According to Theorem 2.6, the incident wavefield u_I is a local WKB function around q_Γ, and within $\mathcal{O}(\epsilon)$ the following holds:

$$u_I(q_\Gamma)=A_\Gamma(q_\Gamma)\exp[iS_\Gamma(q_\Gamma)/\epsilon], \quad (-i\epsilon)\nabla u_I(q_\Gamma)=p_{I,\Gamma}A_\Gamma(q_\Gamma)\exp[iS_\Gamma(q_\Gamma)/\epsilon], \tag{4.11}$$

with

$$A_\Gamma(q_\Gamma)=\mathcal{A}(z_{I,\Gamma})\mathcal{M}(C_{I,\Gamma}),\ S_\Gamma(q_\Gamma)=\mathcal{S}_I(z_{I,\Gamma})+p_{I,\Gamma}q_\Gamma/2. \tag{4.12}$$

Note that by setting the normal and tangent projection operators to

$$\mathcal{T}_\perp = n^\dagger n, \quad \mathcal{T}_\parallel = I-\mathcal{T}_\perp,$$

and the normal impedance ζ_1 in Ω_1 to

$$\zeta_1=\left[s_1^2(q_\Gamma)-|\mathcal{T}_\parallel p_{I,\Gamma}|^2\right]^{1/2},$$

we have

$$\nabla_\Gamma S_\Gamma(q_\Gamma)=\mathcal{T}_\parallel p_{I,\Gamma},\ n_1\zeta_1=\mathcal{T}_\perp p_{I,\Gamma},\ p_{I,\Gamma}=\nabla_\Gamma S_\Gamma(q_\Gamma)+n_1\zeta_1,$$

where n_1 denotes the normal direction of the interface Γ exterior to Ω_1. Suppose that when confined to Γ, the reflected wave u_R and the transmitted wave u_T are expressed as

$$u_R(q_\Gamma)=\mathcal{F}_R A_\Gamma(q_\Gamma)\exp[iS_\Gamma(q_\Gamma)/\epsilon], \quad u_T(q_\Gamma)=\mathcal{F}_T A_\Gamma(q_\Gamma)\exp[iS_\Gamma(q_\Gamma)/\epsilon], \tag{4.13}$$

where $\mathcal{F}_R$ and $\mathcal{F}_T$ are the reflection and transmission coefficients. Then based on the analysis at the beginning of this section, it holds that

$$(-i\epsilon)\nabla u_R(q_\Gamma)=[\nabla_\Gamma S_\Gamma(q_\Gamma)-n_1\zeta_1]\mathcal{F}_R A_\Gamma(q_\Gamma)\exp[iS_\Gamma(q_\Gamma)/\epsilon], \tag{4.14}$$

$$(-i\epsilon)\nabla u_T(q_\Gamma)=[\nabla_\Gamma S_\Gamma(q_\Gamma)+n_1\zeta_2]\mathcal{F}_T A_\Gamma(q_\Gamma)\exp[iS_\Gamma(q_\Gamma)/\epsilon], \tag{4.15}$$

where

$$\zeta_2=\left[s_2^2(q_\Gamma)-|\mathcal{T}_\parallel p_{I,\Gamma}|^2\right]^{1/2}$$

is the normal impedance in Ω_2. Note that when ζ_2 is purely imaginary, the transmitted wave u_T is evanescent along the normal direction n_1.

The wavefield on Ω_1 is a combination of the incident wave u_I and the reflected wave u_R, and the wavefield on Ω_2 consists solely of the transmitted wave u_T. Applying the connection conditions (4.5) and recalling (4.11), (4.13), and (4.14)–(4.15), we derive

$$1+\mathcal{F}_R=\mathcal{F}_T, \quad \zeta_1(1-\mathcal{F}_R)=\zeta_2\mathcal{F}_T,$$

which leads to

$$\mathcal{F}_R=(\zeta_1-\zeta_2)/(\zeta_1+\zeta_2), \quad \mathcal{F}_T=2\zeta_1/(\zeta_1+\zeta_2).$$

For any $\beta\in\{I,R,T\}$, let us set

$$p_{\beta,\Gamma}=\left(\mathcal{T}_{\|}+a_\beta\mathcal{T}_{\perp}\right)p_{I,\Gamma}, \; z_{\beta,\Gamma}=(q_\Gamma,p_{\beta,\Gamma}), \tag{4.16}$$

with

$$a_I=1, \quad a_R=-1, \quad a_T=\zeta_2/\zeta_1.$$

The reflected wave is always propagating. If $\zeta_2>0$, then the transmitted wave is also propagating. In either case, we determine and parameterize the reflected manifold Λ_R and the transmitted manifold Λ_T by

$$\dot{z}_\beta=J\nabla H(z_\beta), \; z_\beta|_{t=\tau}=z_{\beta,\Gamma}.$$

Therefore, the matrix representation $C_\beta=\partial z_\beta/\partial(t,y)$ of the tangent plane satisfies

$$\dot{C}_\beta=J\nabla^2H(z_\beta)C_\beta, \; C_\beta=Q_\beta P_\beta, \; C_\beta|_{t=\tau}=C_{\beta,\Gamma},$$

where

$$C_{\beta,\Gamma}=\partial z_\beta/\partial(t,y)|_{t=\tau}=\left[J\nabla H(z_{\beta,\Gamma})\;\partial_y z_{\beta,\Gamma}-J\nabla H(z_{\beta,\Gamma})\partial_y\tau\right]. \tag{4.17}$$

By recalling (4.12), (4.13), and applying Theorem 2.6, the global geometrical optics approximation to the reflected and transmitted waves is then given by

$$u_\beta(x)=\int_{z_\beta\in\Lambda_\beta}\mathcal{A}_\beta(z_\beta)\exp[i\mathcal{S}_\beta(z_\beta)/\epsilon][\rho_\epsilon(-z_\beta)\phi_\epsilon](x)d\mathrm{vol}, \; \forall\beta\in\{R,T\},$$

where the phase $\mathcal{S}_\beta$ and the amplitude $\mathcal{A}_\beta$ are determined by

$$\dot{\mathcal{S}}_\beta+[z_\beta,\dot{z}_\beta]/2=0, \; \mathcal{S}_\beta|_{t=\tau}=\mathcal{S}_I(z_{I,\Gamma})+p_{I,\Gamma}q_\Gamma/2-p_{\beta,\Gamma}q_\Gamma/2, \tag{4.18}$$

$$\dot{\mathcal{A}}_\beta+\mathrm{tr}\left[Q_\beta^\dagger(iI+J)\nabla^2H(z_\beta)Q_\beta\right]\mathcal{A}_\beta/2=0, \; \mathcal{A}_\beta|_{t=\tau}=\mathcal{F}_\beta\mathcal{G}_\beta\mathcal{A}_I(z_{I,\Gamma}), \tag{4.19}$$

and the geometric factors $\mathcal{G}_\beta$ are defined as

$$\mathcal{G}_\beta=\mathcal{M}(C_{I,\Gamma})/\mathcal{M}(C_{\beta,\Gamma}).$$

The denominator $\mathcal{M}(C_{\beta,\Gamma})$ in the definition of $\mathcal{G}_\beta$ is well-defined if and only if $\mathcal{M}(C_{I,\Gamma})$ is well-defined. Equivalently, $z_{\beta,\Gamma}$ is a regular point of Λ_β if and only if $z_{I,\Gamma}$ is a regular point of Λ_I. Actually, considering $n^\dagger\partial_y q_\Gamma=0$, by (4.17) and (4.16) it holds that

$$C_{\beta,\Gamma,1}\begin{bmatrix}1 & \partial_y\tau\\ 0 & I_{N-1}\end{bmatrix}=\left(\mathcal{T}_{\|}+a_\beta\mathcal{T}_{\perp}\right)[p_{I,\Gamma}\quad \partial_y q_\Gamma]=\left(\mathcal{T}_{\|}+a_\beta\mathcal{T}_{\perp}\right)C_{I,\Gamma,1}\begin{bmatrix}1 & \partial_y\tau\\ 0 & I_{N-1}\end{bmatrix},$$

which yields

$$C_{\beta,\Gamma,1} = \left(\mathcal{T}_{\|} + a_\beta \mathcal{T}_{\perp}\right) C_{I,\Gamma,1}. \tag{4.20}$$

This implies that $C_{\beta,\Gamma,1}$ and $C_{I,\Gamma,1}$ are simultaneously invertible or non-invertible since

$$\det C_{\beta,\Gamma,1} = a_\beta \det C_{I,\Gamma,1}.$$

In the case that $z_{I,\Gamma}$ is singular with respect to the projection map onto the real plane, the incident point q_I is a caustic and the above interface analysis ceases to be valid. However, since

$$\mathcal{G}_\beta^2 = \frac{\mathcal{M}^2(C_{I,\Gamma})}{\mathcal{M}^2(C_{\beta,\Gamma})} = \frac{\det\left(C_{I,\Gamma,1} - iC_{I,\Gamma,2}\right)}{\det\left(C_{\beta,\Gamma,1} - iC_{\beta,\Gamma,2}\right)} \cdot \frac{\det C_{\beta,\Gamma,1}}{\det C_{I,\Gamma,1}} = a_\beta \frac{\det\left(C_{I,\Gamma,1} - iC_{I,\Gamma,2}\right)}{\det\left(C_{\beta,\Gamma,1} - iC_{\beta,\Gamma,2}\right)}, \tag{4.21}$$

we see that $\mathcal{G}_\beta$ as an entirety is always well-defined regardless of whether $z_{I,\Gamma}$ is regular or not. In the following two subsections, we will deduce an expression of $\mathcal{G}_\beta$ which is valid for all incident cases.

4.2. Geometric factors $\mathcal{G}_\beta$ in the case of a planar interface. Analogous to the derivation of (4.17), it is straightforward to verify that

$$C_{I,\Gamma} \begin{bmatrix} 1 & \partial_y \tau \\ 0 & I_{N-1} \end{bmatrix} = \begin{bmatrix} p_{I,\Gamma} & \partial_y q_\Gamma \\ s_1 \nabla s_1 & \partial_y p_{I,\Gamma} \end{bmatrix}.$$

Since $C_{I,\Gamma}$ is a matrix representation of a Lagrangian plane and $\partial_y q_\Gamma$ lies in the tangent plane of Γ at q_Γ, we know that the matrix

$$L = \begin{bmatrix} n & \partial_y q_\Gamma \\ 0 & \mathcal{T}_{\|} \partial_y p_{I,\Gamma} \end{bmatrix}$$

also represents a Lagrangian plane. Set

$$\tilde{P} = \left\{ (\partial_y q_\Gamma)^\dagger \partial_y q_\Gamma + (\partial_y p_{I,\Gamma})^\dagger \mathcal{T}_{\|} \partial_y p_{I,\Gamma} \right\}^{1/2} \in \mathbb{R}^{(N-1)\times(N-1)},$$

and let $L = QP$ be the polar decomposition. Then we have

$$P = \left(L^\dagger L\right)^{1/2} = \begin{bmatrix} 1 & 0 \\ 0 & \tilde{P} \end{bmatrix}, \quad Q = \begin{bmatrix} n & \partial_y q_\Gamma \tilde{P}^{-1} \\ 0 & \mathcal{T}_{\|} \partial_y p_{I,\Gamma} \tilde{P}^{-1} \end{bmatrix}.$$

Note that the matrix

$$V = Q_1 + iQ_2 = \left[n \ \mathcal{T}_{\|} \left(\partial_y q_\Gamma + i \partial_y p_{I,\Gamma}\right) \tilde{P}^{-1} \right] \tag{4.22}$$

is unitary since Q is both orthogonal and Lagrangian.

Now suppose that *the interface* Γ *is planar*. In this case, n, $\mathcal{T}_{\perp}$, and $\mathcal{T}_{\|}$ are constant. Let O be any orthogonal real matrix which transforms the normal direction n to the first Cartesian basis vector, i.e.,

$$On = e_1 = [1, 0, \cdots, 0]^\dagger.$$

For an *arbitrary but fixed* choice of y, let us consider the following canonical transformation in the phase space:

$$z \longrightarrow \tilde{z} = \tilde{z}(z) = R_{V^*}(z - z_c), \ z_c = (q_\Gamma, \mathcal{T}_{\|} p_{I,\Gamma}). \tag{4.23}$$

The associated unitary transformation in the function space is

$$u \longrightarrow \mu_\epsilon(V^*)\rho_\epsilon(z_c)u = \mu_\epsilon((OV)^*)\mu_\epsilon(O)\rho_\epsilon(z_c)u.$$

Since

$$OV = O\left[n\ \mathcal{T}_\parallel(\partial_y q_\Gamma + i\partial_y p_{I,\Gamma})\tilde{P}^{-1}\right] = \left[e_1\ (I - e_1 e_1^\dagger)O(\partial_y q_\Gamma + i\partial_y p_{I,\Gamma})\tilde{P}^{-1}\right],$$

we see that the unitary transformation $\mu_\epsilon((OV)^*)$ keeps the first real variable unchanged. Consequently, $\mu_\epsilon((OV)^*)$ commutes with any linear operator that only acts on the first real variable. In particular, if u satisfies the connection conditions (4.5), then $\mu_\epsilon(V^*)\rho_\epsilon(z_c)u$ satisfies

$$[\mu_\epsilon(V^*)\rho_\epsilon(z_c)u]|_{\tilde{\Gamma},\tilde{\Omega}_2} - [\mu_\epsilon(V^*)\rho_\epsilon(z_c)u]|_{\tilde{\Gamma},\tilde{\Omega}_1} = 0, \tag{4.24}$$

$$e_1^\dagger\nabla[\mu_\epsilon(V^*)\rho_\epsilon(z_c)u]|_{\tilde{\Gamma},\tilde{\Omega}_2} - e_1^\dagger\nabla[\mu_\epsilon(V^*)\rho_\epsilon(z_c)u]|_{\tilde{\Gamma},\tilde{\Omega}_1} = 0, \tag{4.25}$$

where

$$\tilde{\Omega}_1 = O(\Omega_1 - q_\Gamma),\ \tilde{\Omega}_2 = O(\Omega_2 - q_\Gamma),\ \tilde{\Gamma} = O(\Gamma - q_\Gamma) = \{(x_1,\dots,x_N)|x_1 = 0\}.$$

The benefit of the above phase space transformation technique is that in the new frame the manifold $\tilde{z}(\Lambda_\beta)$ for any $\beta \in \{I,R,T\}$ admits a local one-to-one projection to the new real space around $\tilde{z}(z_{\beta,\Gamma})$. This point can be verified in the following manner. Recalling that

$$p_{\beta,\Gamma} = \left(\mathcal{T}_\parallel + a_\beta \mathcal{T}_\perp\right)p_{I,\Gamma},$$

we have

$$\partial_y p_{\beta,\Gamma} = \mathcal{T}_\parallel \partial_y p_{I,\Gamma} + \mathcal{T}_\perp \partial_y(a_\beta p_{I,\Gamma}).$$

The transformed tangent plane $R_{V^*}C_{\beta,\Gamma}$ corresponding to the phase point $z_{\beta,\Gamma}$ satisfies

$$R_{V^*}C_{\beta,\Gamma}\begin{bmatrix} 1 & \partial_y \tau \\ 0 & I_{N-1}\end{bmatrix} = \begin{bmatrix} n^\dagger & 0 \\ \tilde{P}^{-1}(\partial_y q_\Gamma)^\dagger & \tilde{P}^{-1}(\partial_y p_{I,\Gamma})^\dagger \mathcal{T}_\parallel \\ 0 & n^\dagger \\ -\tilde{P}^{-1}(\partial_y p_{I,\Gamma})^\dagger \mathcal{T}_\parallel & \tilde{P}^{-1}(\partial_y q_\Gamma)^\dagger \end{bmatrix}$$

$$\times \begin{bmatrix} \left(\mathcal{T}_\parallel + a_\beta\mathcal{T}_\perp\right)p_{I,\Gamma} & \partial_y q_\Gamma \\ s\nabla s & \mathcal{T}_\parallel\partial_y p_{I,\Gamma} + \mathcal{T}_\perp\partial_y(a_\beta p_{I,\Gamma})\end{bmatrix} = \begin{bmatrix} a_\beta n^\dagger p_{I,\Gamma} & 0 \\ * & \tilde{P} \\ * & * \\ * & * \end{bmatrix}.$$

From the above equations, we see that the upper half part of $R_{V^*}C_{\beta,\Gamma}$ is always nonsingular. Therefore, the new real variable forms a local coordinate of $\tilde{z}(\Lambda_\beta)$ around $\tilde{z}(z_{\beta,\Gamma})$.

Recall that the incident, reflected, and transmitted waves have the following expression:

$$u_\beta(x) = \int_{z_\beta \in \Lambda_\beta} \mathcal{A}_\beta(z_\beta)\exp[i\mathcal{S}_\beta(z_\beta)/\epsilon][\rho_\epsilon(-z_\beta)\phi_\epsilon](x)d\text{vol},\ \forall \beta \in \{I,R,T\}.$$

In the new phase space frame (4.23), we have

$$[\mu_\epsilon(V^*)\rho_\epsilon(z_c)u_\beta](x)$$

$$
\begin{aligned}
&= \int_{z_\beta \in \Lambda_\beta} \mathcal{A}_\beta(z_\beta) \exp[i\mathcal{S}_\beta(z_\beta)/\epsilon] [\mu_\epsilon(V^*)\rho_\epsilon(z_c)\rho_\epsilon(-z_\beta)\phi_\epsilon](x) d\mathrm{vol} \\
&= \int_{z_\beta \in \Lambda_\beta} \mathcal{A}_\beta(z_\beta) \exp\{i[\mathcal{S}_\beta(z_\beta) + [z_c, z_\beta]/2]/\epsilon\} [\rho_\epsilon(-R_{V^*}(z_\beta - z_c))\phi_\epsilon](x) d\mathrm{vol} \\
&= \int_{\tilde{z} = R_{V^*}(z_\beta - z_c) \in \tilde{z}(\Lambda_\beta)} \mathcal{A}_\beta(z_\beta) \exp\{i[\mathcal{S}_\beta(z_\beta) + [z_c, z_\beta]/2]/\epsilon\} [\rho_\epsilon(-\tilde{z})\phi_\epsilon](x) d\mathrm{vol}.
\end{aligned}
$$

These are still extended WKB functions since the transformation $z_\beta \to \tilde{z} = \tilde{z}(z_\beta)$ is canonical. Considering

$$
R_{V^*}(z_{\beta,\Gamma} - z_c) = \begin{bmatrix} n^\dagger & 0 \\ \tilde{P}^{-1}(\partial_y q_\Gamma)^\dagger & \tilde{P}^{-1}(\partial_y p_{I,\Gamma})^\dagger \mathcal{T}_\parallel \\ 0 & n^\dagger \\ -\tilde{P}^{-1}(\partial_y p_{I,\Gamma})^\dagger \mathcal{T}_\parallel & \tilde{P}^{-1}(\partial_y q_\Gamma)^\dagger \end{bmatrix} \begin{bmatrix} 0 \\ a_\beta \mathcal{T}_\perp p_{I,\Gamma} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ a_\beta n^\dagger p_{I,\Gamma} \\ * \end{bmatrix},
$$

we know that at $x = 0$ the following holds within $\mathcal{O}(\epsilon)$

$$
\begin{aligned}
&[\mu_\epsilon(V^*)\rho_\epsilon(z_c)u_\beta](0) = \mathcal{A}_\beta(z_{\beta,\Gamma}) \mathcal{M}(R_{V^*} C_{\beta,\Gamma}) \exp\{i[\mathcal{S}_\beta(z_{\beta,\Gamma}) + [z_c, z_{\beta,\Gamma}]/2]/\epsilon\}, \\
&(-i\epsilon) e_1^\dagger \nabla [\mu_\epsilon(V^*)\rho_\epsilon(z_c)u_\beta](0) = a_\beta n^\dagger p_{I,\Gamma} [\mu_\epsilon(V^*)\rho_\epsilon(z_c)u_\beta](0).
\end{aligned}
$$

Note that since

$$
\mathcal{S}_\beta(z_{\beta,\Gamma}) = \mathcal{S}_I(z_{I,\Gamma}) + (1 - a_\beta) q_\Gamma \mathcal{T}_\perp p_{I,\Gamma}/2,
$$

and

$$
[z_c, z_{\beta,\Gamma}] = [z_c, z_{\beta,\Gamma} - z_c] = [(q_\Gamma, \mathcal{T}_\parallel p_{I,\Gamma}), (0, a_\beta T p_{I,\Gamma})] = a_\beta q_\Gamma \mathcal{T}_\perp p_{I,\Gamma},
$$

we have

$$
[\mu_\epsilon(V^*)\rho_\epsilon(z_c)u_\beta](0) = \mathcal{A}_\beta(z_{\beta,\Gamma}) \mathcal{M}(R_{V^*} C_{\beta,\Gamma}) \exp\{i[\mathcal{S}_I(z_{I,\Gamma}) + q_\Gamma \mathcal{T}_\perp p_{I,\Gamma}/2]/\epsilon\}.
$$

Applying the connection conditions (4.24)–(4.25) yields

$$
\mathcal{A}_\beta(z_{\beta,\Gamma}) \mathcal{M}(R_{V^*} C_{\beta,\Gamma}) = \mathcal{F}_\beta \mathcal{A}_I(z_{I,\Gamma}) \mathcal{M}(R_{V^*} C_{I,\Gamma}), \ \forall \beta \in \{R, T\}.
$$

Recalling

$$
\mathcal{A}_\beta(z_{\beta,\Gamma}) = \mathcal{F}_\beta \mathcal{G}_\beta \mathcal{A}_I(z_{I,\Gamma}), \ \forall \beta \in \{R, T\},
$$

we then derive

$$
\mathcal{G}_\beta = \mathcal{M}(R_{V^*} C_{I,\Gamma}) / \mathcal{M}(R_{V^*} C_{\beta,\Gamma}), \ \forall \beta \in \{R, T\}.
$$

4.3. Geometric factors $\mathcal{G}_\beta$ in the general case. When the interface is not planar, the geometric factors $\mathcal{G}_\beta$ cannot be determined in the way explained in the last subsection. As shown by (4.20), $C_{\beta,\Gamma,1}$ is independent of the curvature information of the interface Γ at q_Γ, thus so is $C_{I,\Gamma,1}$. However, since

$$
C_{\beta,\Gamma,2} \begin{bmatrix} 1 & \partial_y \tau \\ 0 & I_{N-1} \end{bmatrix} = \begin{bmatrix} s \nabla s & \partial_y p_{\beta,\Gamma} \end{bmatrix}, \ \forall \beta \in \{R, T\},
$$

and

$$p_{\beta,\Gamma} = \left(\mathcal{T}_{\|} + a_\beta \mathcal{T}_{\perp}\right) p_{I,\Gamma},$$

by a direct computation, we derive

$$\frac{\partial p_{R,\Gamma}}{\partial y} = (\mathcal{T}_{\|} - \mathcal{T}_{\perp})\partial_y p_{I,\Gamma} - 2(n^\dagger p_{I,\Gamma} + n p_{I,\Gamma}^\dagger)\partial_q n \partial_y q_\Gamma,$$

and

$$\begin{aligned}\frac{\partial p_{T,\Gamma}}{\partial y} = &\left(\mathcal{T}_{\|} + a_T^{-1}\mathcal{T}_{\perp}\right)\partial_y p_{I,\Gamma} + (\zeta_1\zeta_2)^{-1}\mathcal{T}_{\perp} p_{I,\Gamma}\partial_y(s_2^2 - s_1^2)/2 \\ &+ \left[(a_T - 1)n^\dagger p_{I,\Gamma} + \left(a_T^{-1} - 1\right) n p_{I,\Gamma}^\dagger\right]\partial_q n \partial_y q_\Gamma,\end{aligned}$$

which implies that the tangent planes $C_{R,\Gamma}$ and $C_{T,\Gamma}$ indeed depend on the curvature information of Γ.

For an arbitrary but fixed y, let us set

$$\Phi_{y,0}(q) = \Phi(q_\Gamma) + \nabla\Phi(q_\Gamma)\cdot(q - q_\Gamma),$$

and

$$\Phi_{y,\sigma}(q) = (1-\sigma)\Phi_{y,0}(q) + \sigma\Phi(q),\ \forall\sigma\in[0,1].$$

The zero-level set of $\Phi_{y,\sigma}$, denoted by Γ_σ, defines a smooth family of interfaces containing the incident point q_Γ. Let us denote the normal direction of Γ_σ by n_σ and set

$$\mathcal{T}_{\perp,\sigma} = n_\sigma^\dagger n_\sigma, \quad \mathcal{T}_{\|,\sigma} = I - \mathcal{T}_{\perp,\sigma}.$$

Then for any $\sigma\in[0,1]$, it holds that

$$n_\sigma(q_\Gamma) = n(q_\Gamma),\ \mathcal{T}_{\perp,\sigma}(q_\Gamma) = \mathcal{T}_{\perp}(q_\Gamma),\ \mathcal{T}_{\|,\sigma}(q_\Gamma) = \mathcal{T}_{\|}(q_\Gamma),\ \nabla n_\sigma(q_\Gamma) = \sigma\nabla n(q_\Gamma).$$

Since the ray indexed by y penetrates through the interface Γ at q_Γ, so does it for any Γ_σ. The geometric factor $\mathcal{G}_{\beta,\sigma}$ determined by replacing Γ with Γ_σ thus defines a smooth family with respect to $\sigma\in[0,1]$. Let us set

$$\begin{aligned}
C_{R,\Gamma,2}^{(\sigma)} &= \left[s\nabla s\ (\mathcal{T}_{\|} - \mathcal{T}_{\perp})\partial_y p_{I,\Gamma} - 2\sigma(n^\dagger p_{I,\Gamma} + n p_{I,\Gamma}^\dagger)\partial_q n\partial_y q_\Gamma\right]\begin{bmatrix} 1 & -\partial_y\tau \\ 0 & I_{N-1}\end{bmatrix}, \\
C_{T,\Gamma,2}^{(\sigma)} &= \left[s\nabla s \quad \begin{matrix}\left(\mathcal{T}_{\|} + a_T^{-1}\mathcal{T}_{\perp}\right)\partial_y p_{I,\Gamma} + (\zeta_1\zeta_2)^{-1}\mathcal{T}_{\perp} p_{I,\Gamma}\partial_y(s_2^2 - s_1^2)/2 \\ +\sigma\left[(a_T - 1)n^\dagger p_{I,\Gamma} + \left(a_T^{-1} - 1\right) n p_{I,\Gamma}^\dagger\right]\partial_q n\partial_y q_\Gamma\end{matrix}\right]\begin{bmatrix} 1 & -\partial_y\tau \\ 0 & I_{N-1}\end{bmatrix}, \\
C_{\beta,\Gamma}^{(\sigma)} &= \begin{bmatrix} C_{\beta,\Gamma,1} \\ C_{\beta,\Gamma,2}^{(\sigma)}\end{bmatrix},\ \forall\beta\in\{R,T\}.
\end{aligned}$$

Note that we have

$$C_{\beta,\Gamma}^{(1)} = C_{\beta,\Gamma},\ \forall\beta\in\{R,T\}.$$

Recalling (4.21), we have

$$\begin{aligned}\mathcal{G}^2_{\beta,\sigma} &= a_\beta \frac{\det\left(C_{I,\Gamma,1} - iC_{I,\Gamma,2}\right)}{\det\left(C_{\beta,\Gamma,1} - iC^{(\sigma)}_{\beta,\Gamma,2}\right)} \\ &= a_\beta \frac{\det\left(C_{I,\Gamma,1} - iC_{I,\Gamma,2}\right)}{\det\left(C_{\beta,\Gamma,1} - iC^{(0)}_{\beta,\Gamma,2}\right)\det\left[\left(C_{\beta,\Gamma,1} - iC^{(0)}_{\beta,\Gamma,2}\right)^{-1}\left(C_{\beta,\Gamma,1} - iC^{(\sigma)}_{\beta,\Gamma,2}\right)\right]}.\end{aligned}$$

Since $C^{(0)}_{\beta,\Gamma}$ is the matrix representation of the tangent plane when the interface is planar, applying the result in the last subsection, we have

$$\mathcal{G}_{\beta,0} = \mathcal{M}(R_{V^*}C_{I,\Gamma})/\mathcal{M}(R_{V^*}C^{(0)}_{\beta,\Gamma}),\ \forall \beta \in \{R,T\}. \tag{4.26}$$

Therefore, we arrive at

$$\mathcal{G}_{\beta,\sigma} = \mathcal{G}_{\beta,0}\left\{\det\left[\left(C_{\beta,\Gamma,1} - iC^{(0)}_{\beta,\Gamma,2}\right)^{-1}\left(C_{\beta,\Gamma,1} - iC^{(\sigma)}_{\beta,\Gamma,2}\right)\right]\right\}^{-1/2}.$$

In the above equation, the branch of the square root function is chosen to ensure the continuity of $\mathcal{G}_{\beta,\sigma}$ with respect to σ.

As a matter of fact, we can determine $\mathcal{G}_{\beta,\sigma}$, and thus $\mathcal{G}_\beta = \mathcal{G}_{\beta,1}$, without resorting to the above continuity argument. Since $C^{(\sigma)}_{\beta,\Gamma,2}$ is linear with respect to σ, we know that

$$\begin{aligned}W_\sigma &\stackrel{def}{=} \left(C_{\beta,\Gamma,1} - iC^{(0)}_{\beta,\Gamma,2}\right)^{-1}\left(C_{\beta,\Gamma,1} - iC^{(\sigma)}_{\beta,\Gamma,2}\right) \\ &= I + i\left(C_{\beta,\Gamma,1} - iC^{(0)}_{\beta,\Gamma,2}\right)^{-1}\left(C^{(0)}_{\beta,\Gamma,2} - C^{(\sigma)}_{\beta,\Gamma,2}\right) = I - i\sigma\left(C_{\beta,\Gamma,1} - iC^{(0)}_{\beta,\Gamma,2}\right)^{-1}\frac{d}{d\sigma}C^{(\sigma)}_{\beta,\Gamma,2},\end{aligned}$$

where $\frac{d}{d\sigma}C^{(\sigma)}_{\beta,\Gamma,2}$ does not depend on σ. This implies that the eigenvalues of the matrix W_σ would never appear on the negative real axis, the branch cut of the square root function, since, otherwise, there must exist some $\sigma \in [0,1]$ such that W_σ is singular, and this is in contradiction to the assumption that the ray is transversal to the interface at q_Γ. Based on this fact, we have

$$(\det W_\sigma)^{1/2} = \prod_{\lambda \in SpW_\sigma}\sqrt{\lambda} \stackrel{def}{=} \sqrt{\det} W_\sigma, \tag{4.27}$$

where SpW_σ denotes the set of eigenvalues of W_σ (accounting for their multiplicities). Combining (4.26) and (4.27) we finally derive

$$\mathcal{G}_\beta = \frac{\mathcal{G}_{\beta,0}}{\sqrt{\det} W_1} = \frac{\mathcal{M}(R_{V^*}C_{I,\Gamma})}{\mathcal{M}(R_{V^*}C^{(0)}_{\beta,\Gamma})\sqrt{\det}\left[\left(C_{\beta,\Gamma,1} - iC^{(0)}_{\beta,\Gamma,2}\right)^{-1}\left(C_{\beta,\Gamma,1} - iC_{\beta,\Gamma,2}\right)\right]}.$$

5. Miscellaneous issues

In this section, we discuss some issues related to the global geometrical optics approximation and the numerical implementation.

5.1. Phase modification. In the definition of extended WKB functions, we have assumed that the phase $\mathcal{S}$ is a generating function of $pdq-d(pq)/2$. If we modify $\mathcal{S}$ to $\mathcal{S}-pq/2$, then an equivalent definition of the extended WKB function is an asymptotic series of the form

$$u(x)=\int_{z\in\Lambda}[\mathcal{A}(z)+(-i\epsilon)\mathcal{A}_1(z)+\cdots]\exp[i\mathcal{S}(z)/\epsilon]\,[\tilde{\rho}_\epsilon(-z)\phi_\epsilon](x)d\text{vol},\ x\in\mathbb{R}^N,$$

where

$$\tilde{\rho}_\epsilon(-z)=\exp(-ipq/2\epsilon)\rho_\epsilon(-z),$$

and $\mathcal{S}$ is a generating function of pdq of the Lagrangian manifold Λ. Accordingly, the global geometrical optics approximation to the general wave equation (3.1) is an extended WKB function

$$u(x)=\int_{z\in\Lambda}\mathcal{A}(z)\exp[i\mathcal{S}(z)/\epsilon]\,[\tilde{\rho}_\epsilon(-z)\phi_\epsilon](x)d\text{vol},$$

where $\mathcal{S}$ solves the ODE

$$\dot{\mathcal{S}}=p\dot{q}.$$

The new phase function $\mathcal{S}$ has an explicit physical meaning since it is simply the classical action function, also called travel time in geometrical optics. In addition, the ODE (4.18) satisfied by the travel time is modified into a more handy form,

$$\dot{\mathcal{S}}_\beta=p_\beta\dot{q}_\beta,\ \mathcal{S}_\beta|_{t=\tau}=\mathcal{S}_I(z_{I,\Gamma}).$$

This recovers a known fact that the reflected and transmitted waves have the same travel time as the incident wave at the impinging time point.

5.2. Other connection conditions. In Section 3, we specified the C^1-continuity connection conditions at the interface. Actually, the reflected manifold and the transmitted manifold (if they exist) only depend on the slowness fields, and the connection conditions only affect the reflection and/or transmission coefficients. Some other connection conditions can be considered analogously. For example, if the interface is sound-soft or sound-hard, i.e., a homogeneous Dirichlet or Neumann boundary condition is specified, then only the reflected manifold exists and the reflection coefficient $\mathcal{F}_R$ is ±1. If the C^1-continuity connection (4.5) is modified to

$$u|_{\Gamma,\Omega_2}-u|_{\Gamma,\Omega_1}=0,\ \sigma_2 n^\dagger\nabla u|_{\Gamma,\Omega_2}-\sigma_1 n^\dagger\nabla u|_{\Gamma,\Omega_1}=0, \tag{5.1}$$

then the reflection and transmission coefficients are

$$\mathcal{F}_R=\frac{\sigma_1\zeta_1-\sigma_2\zeta_2}{\sigma_1\zeta_1+\sigma_2\zeta_2},\quad \mathcal{F}_T=\frac{2\sigma_1\zeta_1}{\sigma_1\zeta_1+\sigma_2\zeta_2}.$$

The connection conditions (5.1) will appear if one considers the time-harmonic acoustic equation

$$\left[\frac{\omega^2}{\kappa}+\nabla\cdot\left(\frac{1}{\rho}\nabla\right)\right]u=0,$$

where ρ is the density, κ is the bulk modulus, and u is the pressure. If κ is piecewise smooth and ρ is piecewise constant with the interface Γ, then $s=\sqrt{\rho/\kappa}$ and the connection conditions are of the form (5.1) with

$$\sigma_1=\rho^{-1}|_{\Gamma,\Omega_1},\ \sigma_2=\rho^{-1}|_{\Gamma,\Omega_2}.$$

5.3. Conservation property. The ODE (3.13) satisfied by the amplitude function can be rewritten into a more handy form. Put $U = Q_1 + iQ_2$. Then U is unitary. It is straightforward to verify that

$$Q^\dagger (I - iJ)\dot{Q} = \overline{U^* \dot{U}} = \overline{U^{-1}\dot{U}}.$$

From the ODE (3.3), we have

$$\dot{Q}P + Q\dot{P} = J\nabla^2 H(z) QP.$$

Thus,

$$\nabla^2 H(z) Q = J^{-1}\left(\dot{Q}P + Q\dot{P}\right)P^{-1} = J^{-1}\left(\dot{Q} + Q\dot{P}P^{-1}\right).$$

Therefore, we have

$$\begin{aligned} &\operatorname{tr}\left[Q^\dagger (iI + J)\nabla^2 H(z) Q\right] = \operatorname{tr}\left[Q^\dagger (I - iJ)\left(\dot{Q} + Q\dot{P}P^{-1}\right)\right] \\ &= \operatorname{tr}\left[Q^\dagger (I - iJ)\dot{Q} + \dot{P}P^{-1}\right] = \operatorname{tr}\left(\overline{U^{-1}\dot{U}} + \dot{P}P^{-1}\right) = \frac{1}{\det(\bar{U}P)}\frac{d}{dt}\det(\bar{U}P). \end{aligned} \tag{5.2}$$

In the last equality, we have used the Liouville formula; i.e., for any smooth family of nonsingular matrices $\Phi = \Phi(t)$, it holds that

$$\frac{d}{dt}\det\Phi = \operatorname{tr}\left(\Phi^{-1}\frac{d\Phi}{dt}\right)\det\Phi.$$

Substituting (5.2) into the amplitude ODE (3.13) yields

$$\frac{d}{dt}\left[\mathcal{A}^2 \det(\bar{U}P)\right] = \frac{d}{dt}\left[\mathcal{A}^2 \det(C_1 - iC_2)\right] = 0, \tag{5.3}$$

which implies that $\mathcal{A}^2 \det(\bar{U}P)$ is a conserved quantity along each bi-characteristic. Thus the amplitude $\mathcal{A}$ can be determined either from the ODE (3.13) or from a continuity treatment.

5.4. Constant slowness case. When the slowness field is constant, the ODE system determining the global geometrical optics approximation can be integrated out explicitly. In this case, the solution of the Hamiltonian system

$$\dot{q} = p, \quad \dot{p} = 0$$

is simply

$$q(t) = q(t_s) + (t - t_s)p(t_s), \quad p(t) = p(t_s),$$

where t_s is the starting time point. The travel time satisfies

$$\mathcal{S}(t) = \mathcal{S}(t_s) + (t - t_s)|p(t_s)|^2,$$

and the tangent plane satisfies

$$C_1(t) = C_1(t_s) + (t - t_s)C_2(t_s), \quad C_2(t) = C_2(t_s).$$

Applying (5.3) and performing an argument analogous to that in the Subsection 4.3, we derive the amplitude

$$\mathcal{A}(t) = \frac{\mathcal{A}(t_s)}{\sqrt{\det\left([C_1(t_s) - iC_2(t_s)]^{-1}[C_1(t) - iC_2(t)]\right)}}.$$

5.5. Point source. The wave stimulated by a point source is actually an extended WKB function. To illustrate this, let us consider the two-dimensional Helmholtz equation

$$-\epsilon^2 \Delta u - s^2 u = -2i\epsilon^{3/2}\delta(x-x_0).$$

First, we seek a local asymptotic solution around the source point x_0 which admits the following ansatz (see [32, 30])

$$u(x) = \frac{1}{2\epsilon^{1/2}}\left[B(x)+\mathcal{O}(\epsilon)\right] H_0^{(1)}(\Theta/\epsilon). \tag{5.4}$$

Here $H_0^{(1)}$ indicates the zeroth order Hankel function of the first kind. The use of an asymptotic technique analogous to the WKB analysis results in

$$|\nabla\Theta| = s,\ \Theta(x_0) = 0,$$

and

$$2\nabla\Theta\cdot\nabla B + \left(\Delta\Theta - \frac{|\nabla\Theta|^2}{\Theta}\right) B = 0,\ B(x_0) = 1. \tag{5.5}$$

Away from the source point x_0, we have the local asymptotic approximation

$$u(x) \sim \frac{B(x)}{2\epsilon^{1/2}}\left(\frac{2}{\pi\Theta/\epsilon}\right)^{1/2} \exp\left[i(\Theta/\epsilon - \pi/4)\right] = B(x)(2\pi i\Theta)^{-1/2}\exp\left(i\Theta/\epsilon\right).$$

Let us set $s_0 = s(x_0)$ and define the isotropic manifold

$$\Lambda_0 = \{z_0 = (q_0, p_0) | q_0 = x_0, p_0 = s_0[\cos y; \sin y], \forall y \in \mathbb{R}/\mathbb{Z}\}.$$

We determine a Lagrangian manifold Λ_+ by solving the Hamiltonian system

$$\dot{q} = p,\ q|_{t=0} = q_0,$$
$$\dot{p} = s\nabla s,\ p|_{t=0} = p_0,$$

with $t > 0$, and the travel time by

$$\dot{\mathcal{S}} = p\dot{q},\ \mathcal{S}|_{t=0} = 0.$$

Note that Λ_+ is a manifold with boundary Λ_0. Since $\mathcal{S}$ is the generating function of pdq on Λ_+ and Λ_+ admits a local one-to-one projection to the real plane, we have

$$\mathcal{S}(z) = \Theta(q).$$

Let us consider the extended WKB function

$$\tilde{u}(x) = \int_{z\in\Lambda_+} \mathcal{A}(z)\exp\left[i\mathcal{S}(z)/\epsilon\right]\left[\tilde{\rho}_\epsilon(-z)\phi_\epsilon\right](x)d\mathrm{vol}, \tag{5.6}$$

where the amplitude function $\mathcal{A}$ is determined by

$$\dot{C} = J\nabla^2 H(z)C,\ C = QP,\ C|_{t=0} = \left[J\nabla H(z_0)\ \partial_y z_0\right],$$
$$\dot{\mathcal{A}} + \mathrm{tr}\left[Q^\dagger(iI+J)\nabla^2 H(z)Q\right]\mathcal{A}/2 = 0,\ \mathcal{A}|_{t=0} = (2\pi)^{-1/2}s_0^{-1}.$$

We will show that away from x_0, within $\mathcal{O}(\epsilon)$, the function $\tilde{u}$ in (5.6) is the same as the function u given in (5.4).

Away from the source point x_0, $\tilde{u}$ solves the homogeneous Helmholtz equation. Besides, locally within $\mathcal{O}(\epsilon)$, it holds that

$$\tilde{u}(q) = \mathcal{A}(z)\mathcal{M}(C(z))\exp[i\mathcal{S}(z)/\epsilon].$$

Let us set

$$\tilde{B}(q) = \mathcal{A}(z)\mathcal{M}(C(z))[2\pi i\Theta(q)]^{1/2}.$$

It suffices to show that $\tilde{B}$ solves Equation (5.5). First, it is straightforward to verify that

$$\tilde{B}(q_0) = \lim_{t\to 0^+}\mathcal{A}(z)\mathcal{M}(C(z))[2\pi i\Theta(q)]^{1/2} = \lim_{t\to 0^+}\left\{\mathcal{A}|_{t=0}(it)^{-1/2}[2\pi i t s_0^2]^{1/2}\right\} = 1.$$

Second, by Liouville's lemma, we have

$$\frac{d}{dt}\det C_1 = \mathrm{tr}(\dot{C}_1 C_1^{-1})\det C_1 = \mathrm{tr}(C_2 C_1^{-1})\det C_1 = \Delta\Theta \det C_1. \tag{5.7}$$

Since

$$\frac{\Theta}{\tilde{B}^2} = \frac{1}{2\pi i\mathcal{A}^2\mathcal{M}^2} = \frac{\det C_1}{2\pi i\mathcal{A}^2\det(C_1 - iC_2)},$$

recalling (5.3) and applying (5.7), we derive

$$\frac{d}{dt}\frac{\Theta}{\tilde{B}^2} = \Delta\Theta\frac{\Theta}{\tilde{B}^2},$$

which leads to

$$\nabla\Theta\cdot\nabla\frac{\Theta}{\tilde{B}^2} = \Delta\Theta\frac{\Theta}{\tilde{B}^2}.$$

It is straightforward to verify that the above equation is equivalent to Equation (5.5).

6. Numerical experiments

The global geometrical optics approximation (3.14) to the high frequency wave equations is an integral on the underlying Lagrangian manifold. In the numerical implementation, this integral should be discretized with some quadrature scheme. In the (t,y) coordinate, the approximate wavefield u is expressed as

$$u(x) = \int \mathcal{A}(z)\exp[i\mathcal{S}(z)/\epsilon][\rho_\epsilon(-z)\phi_\epsilon](x)\det P dt dy.$$

This integral is then approximated by the trapezoid rule. As suggested by [24, 20, 31] and validated by our partial numerical evidences, the sampling resolution needs only to be $\mathcal{O}(\epsilon^{1/2})$ to ensure optimal first order asymptotic accuracy. In all the numerical tests shown below, we simply set the sampling resolution sufficiently refined.

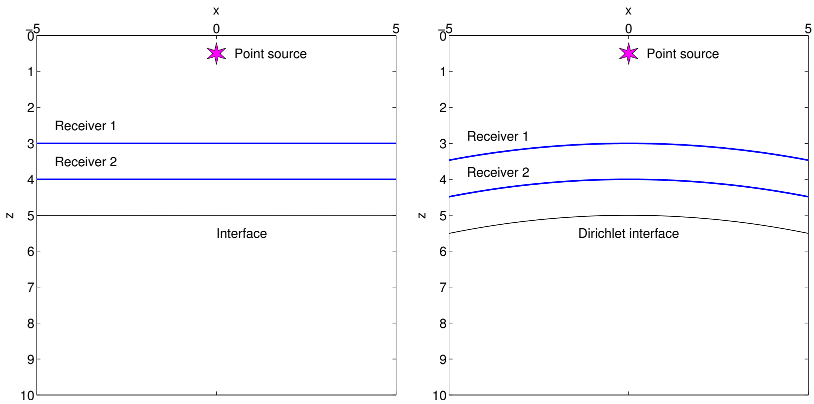

FIG. 6.1. *Dirichlet interface. The source point is located at* $(0,0.5)$*. Left: a line with* $z=5$*. Right: a circle of radius 25 with the center located at* $(0,30)$*.*

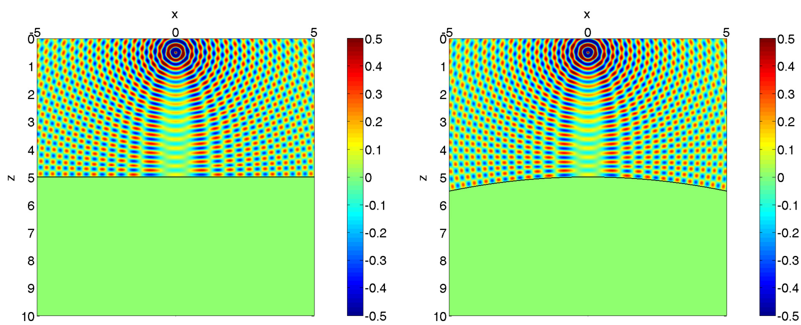

FIG. 6.2. *Green function with a Dirichlet boundary. The angular frequency is* $\omega=20$*.*

6.1. Sound-soft Green function. In this part, we evaluate the numerical asymptotic accuracy of the global geometrical optics approximation to the Green function of the Helmholtz equation with a homogeneous Dirichlet boundary. The point source is located at $(0,0.5)$ and the slowness is set to 1. Two cases will be considered. The interface of the first case is a line with $z=5$, and the interface of the second one is a circle of radius 25 with the center located at $(0,30)$. See Figure 6.1.

In Figure 6.2, we show the Green function approximated by the global geometrical optics method with an angular frequency of $\omega=20$. In Figure 6.3, we plot the relative maximum errors of the global geometrical optics approximation. For the first example, the exact Green function is available by the mirror principle. For the second example, we compute a reference solution by the separation of variables on the circular boundary. In either case, a first order asymptotic accuracy can be observed.

6.2. Green function for multi-layered media. In this part, we consider the application of the global geometrical optics approximation to the Green function of the Helmholtz equation with multilayered media. The point source is located at $(0,0.5)$, and the computational domain is restricted to $[-5,5]\times[0,10]$. In the first numerical example, we consider a three-layered piecewise constant medium. The velocity on each layer is

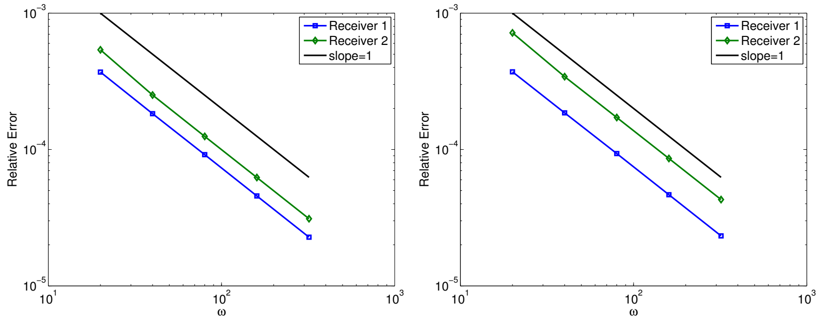

FIG. 6.3. *Relative maximum errors by the global geometrical optics approximation.*

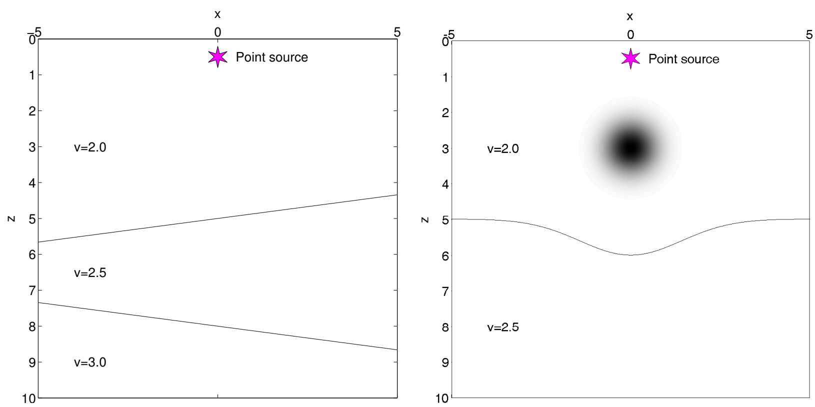

FIG. 6.4. *Multi-layered media. Left: three layers with constant slowness. Right: two layers with variable slowness.*

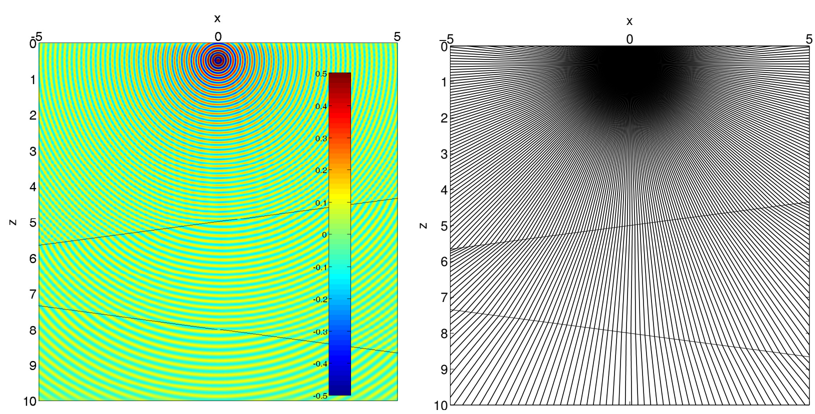

FIG. 6.5. *Three-layered medium. Left: wavefield with frequency 15Hz. Right: incident and first transmitted rays.*

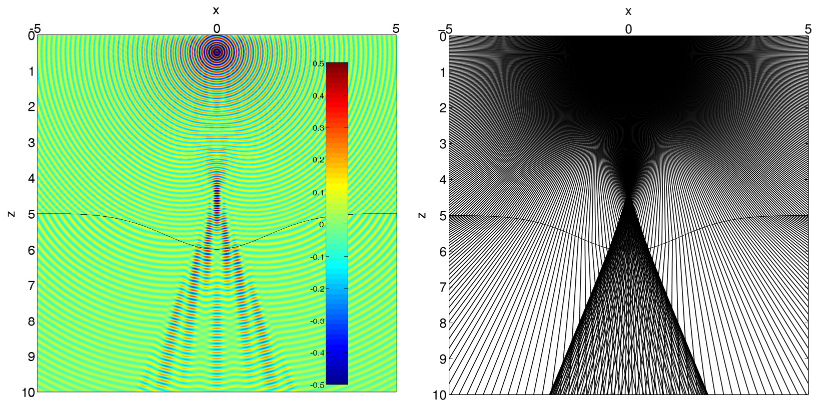

FIG. 6.6. *Two-layered medium with variable slowness. Left: wavefield with frequency 15Hz. Right: incident and transmitted rays.*

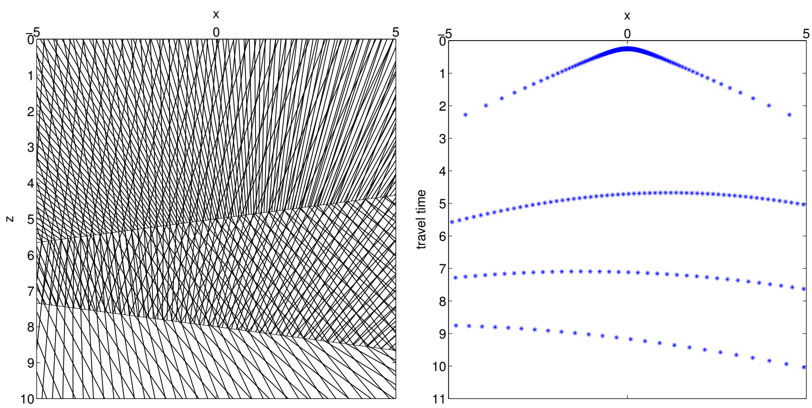

FIG. 6.7. *Three-layered medium. Left: reflected and transmitted rays. Right: travel times for the rays traveling to* $z = 0$.

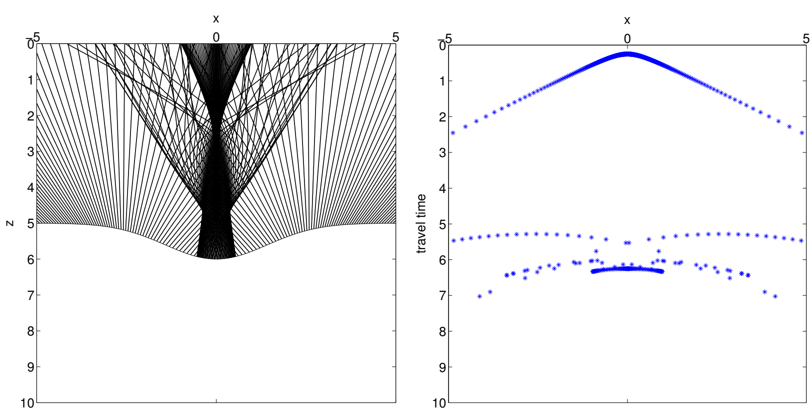

FIG. 6.8. *Two-layered medium with variable slowness. Left: reflected rays. Right: travel times for the rays traveling to* $z = 0$.

2.0, 2.5, and 3.0 respectively. The two interfaces pass through the points $(0,5)$ and $(0,8)$ and have a dip angle of 7.5 degrees. In the second numerical example, we consider a two-layered medium with variable slowness. The interface is a curve determined by

$$z = 5 + \exp(-x^2/4).$$

The slowness in the upper part is

$$s(x,z) = 0.5 + 0.2\exp\left\{-[x^2 + (z-3)^2]/0.5\right\},$$

and the slowness in the lower part is $1/2.5 = 0.4$. See Figure 6.4.

On the left of Figure 6.5 and Figure 6.6, we show the real part of the wavefield derived by global geometrical optics approximation with frequency 15 Hz, and on the right, we show the incident rays and the first transmitted rays. Since the reflected waves are relatively weak, the wave pattern is in good agreement with the wavefronts of the incident and first transmitted waves. On the left of Figure 6.7 and Figure 6.8, we show the reflected and successive transmitted rays, and on the right, we plot the travel times for all rays which travel to the line $z = 0$.

7. Conclusion

The global geometrical optics approximation is a new semi-classical approximation of high frequency wave equations which was originally proposed in [33]. In this paper, we rederived it in a more concise way. It was shown that the right ansatz for the asymptotic approximation of the high frequency waves is the extended WKB function, as opposed to the WKB function used in the classical geometrical optics approximation. The new ingredient of this paper is an interface analysis of the Helmholtz equation when the incident waves are extended WKB functions. Though this analysis is standard for the regular incident points, it becomes more involved when the incident points are singular. The key issue is to determine the geometric factor which relates to the incident and reflected (or transmitted) manifolds. Instead of resorting to the continuity treatment, we developed a method which only utilizes the local properties of the incident rays and the interface geometry.

REFERENCES

[1] G. Ariel, B. Enquist, N.M. Tanushev, and R. Tsai, *Gaussian beam decomposition of high frequency wave fields using expectation-maximization*, J. Comput. Phys., 230(6), 2303–2321, 2011.

[2] V.M. Babič and V.S. Buldyrev, *Short-wavelength Diffraction Theory: Asymptotic Methods*, Springer Series on Wave Phenomena, Springer-Verlag, 4, 1991.

[3] I.M. Babuska and S.A. Sauter, *Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers*, SIAM Review, 42(3), 451–484, 2000.

[4] B. Engquist and L. Ying, *Sweeping preconditioner for the Helmholtz equation: hierarchical matrix representation*, Commun. Pure Appl. Math., 64, 697–735, 2011.

[5] B. Engquist and L. Ying, *Sweeping preconditioner for the Helmholtz equation: moving perfectly matched layers*, Multiscale Model. Simul., 9, 686–710, 2011.

[6] O.G. Ernst and M.J. Gander, *Why It Is Difficult To Solve Helmholtz Problems With Classical Iterative Methods*, Numerical Analysis of Multiscale Problems, I.G. Graham et al. (eds.), Springer-Verlag, Berlin, Heidelberg, 325–363, 2012.

[7] G.B. Folland, *Harmonic Analysis in Phase Space*, Princeton University Press, 1989.

[8] E.J. Heller, *Time-dependent approach to semiclassical dynamics*, J. Chem. Phys., 62(4), 1544–1555, 1975.

[9] M.F. Herman and E. Kluk, *A semiclassical justification for the use of non-spreading wavepackets in dynamics calculations*, Chem. Phys., 91, 27–34, 1984.

[10] S. Jin, H. Wu, and X. Yang, *Gaussian beam methods for the Schrödinger equation in the semi-classical regime: Lagrangian and Eulerian formulations*, Commun. Math. Sci., 6, 995–1020, 2008.

[11] S. Jin, H. Wu, and X. Yang, *Semi-Eulerian and high order gaussian beam methods for the Schrodinger equation in the semiclassical regime*, Commun. Comp. Phys., 9, 668–687, 2011.

[12] K. Kay, *The Herman–Kluk approximation: derivation and semiclassical corrections*, Chem. Phys., 322, 3–12, 2006.

[13] Y.A. Kravtsov and Y.A. Orlov, *Geometrical Optics of Inhomogeneous Media*, Springer-Verlag, Berlin, 1990.

[14] H. Liu and J. Ralston, *Recovery of high frequency wavefields for the acoustic wave equation*, Multiscale Model. Sim., 8(2), 428–444, 2009.

[15] H. Liu and J. Ralston, *Recovery of high frequency wavefields from phase space-based measurements*, Multiscale Model. Sim., 8(2), 622–644, 2010.

[16] H. Liu, O. Runborg, and N.M. Tanushev, *Error estimates for Gaussian beam superpositions*, Math. Comp. 82, 919–952, 2013.

[17] J. Lu and X. Yang, *Frozen Gaussian approximation for general linear strictly hyperbolic systems: formulations and Eulerian methods*, Multiscale Model. Sim., 10, 451–472, 2012.

[18] J. Lu and X. Yang, *Convergence of frozen Gaussian approximation for high frequency wave propagation*, Commun. Pure Appl. Math., 65, 759–789, 2012.

[19] P.V. Maslov and M.V. Fedoryuk, *Semi-classical Approximation in Quantum Mechanics*, Dordrecht Reidel Pub. Co., 1981.

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

[21] V. Rousse, *Semiclassical simple initial value representations*, Université Paris 12, arXiv: 0904.0387, 2009.

[22] H. Sato and M.C. Fehler, *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*, New York, Springer, 1989.

[23] A.H. Sheikh, D. Lahaye, and C. Vuik, *On the convergence of shifted Laplace preconditioner combined with multilevel deflation*, Numer. Linear Alg. Appl., 20(4), 645–662, 2013.

[24] N.M. Tanushev, *Superpositions and higher order Gaussian beams*, Commun. Math. Sci., 6, 449–475, 2008.

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

[26] F.D. Tappert, *Diffractive ray tracing of laser beams*, J. Opt. Soc. Am., 66(12), 1368–1373, 1976.

[27] F.D. Tappert, *Wave Propagation and Underwater Acoustics*, Lecture Notes in Physics, Berlin: Springer, 70, 1977.

[28] H. Wu, Z. Huang, S. Jin, and D. Yin, *Gaussian beam methods for the Dirac equation in the semi-classical regime*, Commun. Math. Sci., 10, 1301–1315, 2012.

[29] J. Xia, S. Chandrasekaran, M. Gu, and X.S. Li, *Superfast multifrontal method for large structured linear systems of equations*, SIAM J. Matrix Anal. Appl., 31(3), 1382–1411, 2009.

[30] M.J. Yedlin, *Uniform asymptotic solution for the Green function for the two-dimensional acoustic equation*, J. Acoust. Soc. Am., 81(2), 238–243, 1987.

[31] D. Yin and C. Zheng, *Composite coherent states approximation for one-dimensional multi-phased wave functions*, Commun. Comput. Phys., 11(3), 951–984, 2012.

[32] E. Zauderer, *Uniform asymptotic solutions of the reduced wave equation*, J. Math. Anal. Appl., 30, 157–171, 1971.

[33] C. Zheng, *Global geometrical optics method*, Commun. Math. Sci., 11(1), 105–140, 2013.