COMMUN. MATH. SCI.
Vol. 11, No. 1, pp. 105–140
© 2013 International Press

GLOBAL GEOMETRICAL OPTICS METHOD*

CHUNXIONG ZHENG†

Abstract. We develop a novel approach, named the *global geometrical optics method*, for the numerical solution to wave equations in the high-frequency regime. The initial Cauchy data is assumed to be in the WKB form. We first study the Schrödinger equation, and then extend relevant results to the general scalar wave equations. The basic idea of this approach is to reformulate the governing equation in a moving frame, and to derive a WKB-type function merely defined on the Lagrangian manifold induced by the Hamiltonian flow. From this WKB-type function, the wave solution can be retrieved to within first order accuracy by a coherent state integral. The merit of the proposed approach is manyfold. Firstly, compared with the thawed Gaussian beam approaches, it presents an approximate wave solution with first order asymptotic accuracy pointwise, even around caustics. Secondly, compared with the canonical operator method, this approach does not require any a priori knowledge about the structure of the Lagrangian manifold. Thirdly, compared with the frozen Gaussian beam approaches such as the Herman-Kluk semi-classical propagator method, the proposed approach involves an integral on a manifold of much lower dimension. We report numerical tests on both Schrödinger and Helmholtz equations.

Key words. Hamiltonian system, coherent state, unitary representation, Lagrangian manifold, caustics, semi-classical approximation.

AMS subject classifications. 65M25, 35F10, 78M35.

1. Introduction

Wave propagation is a fundamental issue of study in various disciplines. The main ingredients of this study include the identification of wave patterns, the mechanism of wave interactions, the simulation of wave phenomena, and so on. Though for a long time much attention has been attracted to the nonlinear wave behaviors, the study of linear waves still remains a hot research subject, especially when the medium is heterogeneous and the wave oscillates on a rather small length scale. In the latter case, we say the wave propagation bears a high-frequency characteristic.

The high-frequency characteristic presents a great challenge to the scientists in any related research area. The difficulty is obvious from the computational point of view, since generally one cannot afford to numerically resolve the highly oscillatory wave field with grid points. This situation becomes much worse, and even hopeless in many dimensions. Therefore, efficient numerical methods and approximate models based on high-frequency asymptotics are significantly motivated.

The geometrical optics method is a classical approach to handle the high-frequency characteristic. This method attempts to seek a wave solution u in the so-called *WKB* (Wentzel-Kramers-Brillouin; see [22]) form:

$$u(x) = A^\epsilon(x)\exp\left(\frac{iS(x)}{\epsilon}\right),\ A^\epsilon = \sum_{j=0}^{\infty}(-i\epsilon)^j A_j. \tag{1.1}$$

Here ϵ is the reference small wavelength, S is a smooth real function called the *phase*, and the A_j are smooth functions (generally complex-valued) called *amplitudes*. By inserting (1.1) into the governing wave equation (scalar case), say,

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

*Received: April 29, 2012; accepted (in revised form): July 5, 2012. Communicated by Shi Jin.
†Department of Mathematical Sciences, Tsinghua University, Beijing 100084, P.R. China (czheng@math.tsinghua.edu.cn).

and equating the different powers of ϵ, one derives

- a Hamilton-Jacobi equation for the phase S:

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

- a sequence of *transport equations* for the amplitudes A_j, which can be solved successively after the phase S is uniquely determined.

The solution (if it exists) of the Hamilton-Jacobi equation (1.3), which is compatible with prescribed boundary data, can be computed by the method of characteristics. However, this solution is not ensured to be valid in a sufficiently large domain which encloses the region of physical interest; *caustics*, where some rays intersect, will generally come into being. In this case, the classical geometrical optics method fails to present a global asymptotic approximation.

Actually, the caustic problem stems from the rigid choice of position representation for the wave solution. While projected to the real space, the Lagrangian manifold displaced by the Hamiltonian flow becomes multi-valued around the caustic points. It is known that the phase function admits a natural continuation to the Lagrangian manifold by performing suitable line integral along trajectories. When singularities appear while projecting the local Lagrangian manifold to the real space, one can still derive a single-valued phase function by choosing another suitable representation. This observation motivated Maslov [28] to develop the canonical operator method. The basic idea consists of three steps. First, the whole Lagrangian manifold is decomposed into a union of small pieces, such that each piece admits at lease one single-valued mixed position-momentum representation. Second, a local formal asymptotic approximation is derived by the classical geometrical optics method. In the end, all local asymptotic approximations are patched together to form a global approximation with the aid of *Morse-Maslov index*.

Maslov's theory is mathematically beautiful, but it does not end the story of linear high-frequency waves. A practical but inevitable issue in Maslov's approach is how to set up a specific partition of unity which fulfills the underlying requirements. This is of course case-dependent, and to solve it one needs to understand in advance the structure of the associated Lagrangian manifold. However, this task is by no means a trivial matter in many 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 semi-classical approximation method which directly resolves the high-frequency wave field. This approach was initiated in the 1960s (see [1]), and later applied to the study of singularity propagation by Hörmander [14] and Ralston [31]. Historically, this approach is divided into two subclasses—thawed-type (see [12]) and frozen-type (see [13])—depending on whether the beam width is tunable or not during the time evolution. A common point shared by these two subclasses is that the central curve of each Gaussian beam is exactly a specific classical ray obeying the Hamiltonian system. The difference lies in their methodology of function approximation. In the thawed-type approach, each individual beam is an asymptotic solution of the governing wave equation. This renders the beam summation a suitable asymptotic solution in a large domain. Comparatively, in the frozen-type approach, each individual beam is not an asymptotic solution. The accuracy of the beam summation is realized 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 [33, 21], and elegant from the methodological point of view. In recent years, many papers have addressed the accuracy analysis of this method; see [29, 25, 26, 27]. The Gaussian beam approach has also been applied to boundary value problems of high-frequency waves [4, 34].

If one is not interested in the solution behavior around caustics, there are some other choices, especially more efficient in lower dimensions, to retrieve the multi-phased wave information. It has been shown by Gerard et al [9] that the Wigner transform of the high-frequency waves converges weakly to the measure-valued solution of the Liouville equation. The moment closure method based on this fact was first introduced by Brenier and Corrias [3] for multi-branch entropy solutions of scalar conservation laws. Later it was used by Engquist and Runborg [7] for the computation of multi-phased solutions to the wave equation, and by Gosse [10] for the one-dimensional Schrödinger equation. Jin and his collaborators [11, 16] systematically developed this method for the semi-classical Schrödinger equation and some related problems. The moment method offers great efficiency compared with the phase space computation since it leads to an Eulerian method merely formulated in the physical space. However, this method becomes much more complicated (see [15]) if there exist many phase branches or in high dimensions.

A more elegant way to retrieve the multi-phased information is the level set method, pioneered by Osher and his collaborators. In [6] Cheng, Liu, and Osher applied it for the semi-classical Schrödinger equations. Jin and Osher [18] presented the general framework of the level set method for computing multi-phased solutions to quasi-linear hyperbolic PDEs and the Hamilton-Jacobi equation. Liu, Cheng, and Osher studied the general framework for any first order PDE in [23]. It was realized in [17, 24] that the level set method can resolve not only the phase information, but also physical observables. Recently, Jin and Yang [20] improved this method by taking into account the phase shift based on the *a priori* information about the Lagrangian manifold.

In this paper, we are concerned with the solution to the general scalar wave equation

$$H(x,-i\epsilon\nabla)u=0,\ x\in\mathbb{R}^N. \tag{1.4}$$

Here ϵ is a small parameter and H is quantized by Weyl's method. We suppose H is sufficiently smooth, and the Hamiltonian flow associated with H exists globally. The initial Cauchy data is assumed in the WKB form. We will propose a new semi-classical approximation to the wave solution within first order accuracy simply by the method of characteristics. As will be seen, the new approach also resorts to the WKB analysis, just as the classical geometrical optics method does. Moreover, our method is completely blind to the structure of the associated Lagrangian manifold: the location and strength of caustics will automatically manifest after the wave field is constructed through a coherent state integral. Besides, there is no locality constraint on the availability of the new approach. Considering these points, we name the new approach the *global geometrical optics method*. The key ingredient to make these nice properties available is a proper application of a moving frame technique.

The rest of this paper is organized as follows. In Section 2 we collect and derive some fundamental materials necessary for the later analysis. The book [8] by Folland is heavily consulted. The most subtle point in this part is a single-valued unitary representation of the maximum compact symplectic subgroup $Sp(N)\cap O(2N)$, which is equivalent to the unitary matrix group $\mathcal{U}(N)$. It is this unitary representation which makes the global semi-classical approximation available.

In Section 3, we show that there exists a natural connection between the WKB functions expressed in the position coordinate and the WKB-type functions defined on the Lagrangian manifold. Loosely speaking, any WKB function in the position coordinate corresponds to a WKB-type function on the associated Lagrangian manifold. On the other hand, if the phase of WKB-type function satisfies some compatibility condition with the underlying Lagrangian manifold, a WKB function with the same Lagrangian manifold can be formed through a coherent state integral. This connection suggests that it would be more natural to understand (or continue) the WKB function with some WKB-type function on the Lagrangian manifold, since it is the Lagrangian manifold which is the most fundamental object for a specific Cauchy problem.

In Section 4, we study the global asymptotic approximation for the Schrödinger equation, which is a specific example of (1.4) if we combine the spatial and temporal variables together. We deduce an equation satisfied by the wave function expressed in a moving frame. The most remarkable gain with this change of representation is that the new equation admits a global WKB approximation if the initial Lagrangian sub-manifold is sufficiently simple (the exact meaning of this simpleness will be clear to the readers after Section 2). Besides, the phase and the amplitude at the specific spatial point $x=0$ are determined by two ODEs. Changing the moving frame, we derive a WKB-type function on the entire Lagrangian manifold. A first order asymptotic approximation is then derived through a coherent state integral on the simple Lagrangian manifold. In the general case, we employ the idea of partition of unity. A global first order approximation is finally obtained by simply replacing the integrating Lagrangian manifold. A key point here is that this global approximation actually does not rely on the specific choice of partition of unity. Compared with Maslov's method, this implies that the proposed approach is blind to the structure of Lagrangian manifold.

In Section 5, we extend the results of Section 4 to the general scalar wave equations. We build a quantization operator, an analogy to Maslov's canonical operator, as a coherent state integral on the Lagrangian manifold. If the amplitude is determined by a suitable form of ODE, this quantization operator presents a global first order asymptotic approximation to the exact wave solution. Some numerical tests will be reported in Section 6, and Section 7 concludes this paper.

2. Preliminaries

We indicate I as the identity mapping in a vector space. We do not distinguish the column and row forms for any k-tuple $x=(x_1,\cdots,x_k)$ of numbers, functions or operators. We denote by xy the first order contraction for any two tensor objects with appropriate dimensions. For example, if A denotes a linear mapping and x, y are two vector objects, we define

$$xAy=yA^\dagger x=x\cdot Ay=\sum x_jA_{jk}y_k,\ x^2=xx.$$

As usual, $(\cdot)^\dagger$ indicates the transpose operator. When confusion may occur, we use the first order contraction operator $(\cdot)$ explicitly. The second and third order contraction operators are denoted by $(:)$ and $(\therefore)$ respectively. We use the notation $(\otimes)$ to indicate the tensor product of two vector objects. In particular, we set

$$x^{\otimes 2}=x\otimes x,\ x^{\otimes 3}=x\otimes x\otimes x.$$

2.1. Lagrangian plane and simple manifold. We denote by $\mathbb{R}_D^{2N\times D}$ the subset of $\mathbb{R}^{2N\times D}$ with maximum rank D. Given a matrix in $\mathbb{R}_D^{2N\times D}$, its D column vectors span a D-plane (a subspace of dimension D) of $\mathbb{R}^{2N}$. On the other hand, any D-plane of $\mathbb{R}^{2N}$ can be spanned by a matrix in $\mathbb{R}_D^{2N\times D}$. In this sense, we represent all D-planes of $\mathbb{R}^{2N}$ with $\mathbb{R}_D^{2N\times D}$. Of course, there exist many representations for a specific D-plane: if C is one of the choices, then CV represents the same D-plane for any nonsingular matrix $V\in\mathbb{R}^{D\times D}$. We say CV is an equivalent representation of C.

Any D-plane C has an orthogonal representation, i.e., a representation with orthogonal matrix. Let $C=QP$ be the polar decomposition, where $P=(C^*C)^{\frac{1}{2}}$ and Q is orthogonal. Then Q represents the same D-plane as C. It should be noted that the orthogonal representation of a D-plane is also not unique, since Q and QQ' represent the same D-plane for any orthogonal matrix $Q'\in\mathbb{R}^{D\times D}$. A D-plane C is called *Lagrangian* if $C^*JC=0$ (see page 102 of [28]), where $J=\begin{pmatrix} 0 & I \\ -I & 0\end{pmatrix}$ is the standard symplectic matrix. This definition is proper since any other equivalent representation of C, say $\tilde{C}=CV$ with $V\in\mathbb{R}^{D\times D}$ nonsingular, satisfies

$$\tilde{C}^*J\tilde{C}=V^*C^*JCV=0.$$

Any Lagrangian N-plane C admits a representation by a unitary matrix. Let $C=QP$ be the polar decomposition with $Q=\begin{pmatrix} Q_1 \\ Q_2\end{pmatrix}$. Then $U=Q_1+iQ_2$ is unitary. On the other hand, for any unitary matrix $U=U_R+iU_I$, the N-plane $\begin{pmatrix} U_R \\ U_I\end{pmatrix}$ is Lagrangian. Therefore, we can represent all Lagrangian N-planes with $\mathcal{U}(N)$, the unitary transformation group in $\mathbb{C}^N$. Again, this representation is not unique, since U and UQ represent the same Lagrangian plane for any *real* orthogonal matrix $Q\in\mathbb{R}^{N\times N}$.

A Lagrangian manifold Λ in the phase space $\mathbb{R}^{2N}$ is a smooth manifold such that the differential 2-form $dq\wedge dp$ vanishes on Λ (see page 8 of [28]). An equivalent description is such that the tangent plane at each point of Λ is Lagrangian. Given a WKB function $\varphi(x)=A(x)\exp\left(\frac{iS(x)}{\epsilon}\right)$, it is easy to show that the manifold $\Lambda=\{(q,\nabla S(q))|q\in\mathrm{supp}A\}$ is Lagrangian. Here $\mathrm{supp}(A)$ denotes the smallest closed simply-connected domain which encloses all non-zero points of A. We say Λ is the Lagrangian manifold associated with the WKB function φ. For each phase point $z=(q,p)\in\Lambda$, the tangent N-plane at z is $\begin{pmatrix} I \\ \nabla^2 S(q)\end{pmatrix}$, thus it has a unitary matrix representation of the form

$$U=[I+(\nabla^2S(q))^2]^{-\frac{1}{2}}(I+i\nabla^2S(q)).$$

Definition 2.1. *A manifold $\Lambda\subset\mathbb{R}^{2N}$ is called* simple *if Λ admits a diffeomorphic projection on the tangent plane for each $z\in\Lambda$. A WKB function $\varphi(x)=A(x)\exp\left(\frac{iS(x)}{\epsilon}\right)$ is called simple if its associated Lagrangian manifold Λ is simple. Given a Hamiltonian $H(q,p,t)$, φ is called simple up to time $T>0$ if the displaced Lagrangian manifold $\Lambda(t)=g^t\Lambda$ is simple for any $t\in[0,T]$. Here g^t denotes the displacement operator on the phase space $\mathbb{R}^{2N}$ induced by the Hamiltonian H.*

To clarify the concept of simple, let us consider an arc $\Lambda=\{(\cos\theta,\sin\theta)|\theta\in[0,\theta_m]\}$ in $\mathbb{R}^2$. Here, $\theta_m\in(0,2\pi)$. The tangent plane of each point on Λ is actually a straight line. If $\theta_m<\pi/2$, different points on Λ have different projections onto an arbitrary tangent line. Therefore, the manifold Λ with $\theta_m<\pi/2$ is simple according to our definition. However, if $\theta_m>\pi/2$, the manifold Λ is not simple, since the points with angle $\pi/2\pm\delta$ have the same projection onto the tangent line of point (1,0). Here, δ is chosen such that $\delta\in(0,\pi/2)$ and $\pi/2+\delta\le\theta_m$.

Throughout this paper, we assume that the Hamiltonian flow exists globally, i.e., g^t is well-defined for any $t\in\mathbb{R}$. Suppose Λ is the Lagrangian manifold associated with $\varphi(x)=A(x)\exp\left(\frac{iS(x)}{\epsilon}\right)$. By a simple compactness argument, we know that given an arbitrary but finite $T>0$, there exists a constant $\sigma_T>0$ such that for any partition of unity $\{e_j(x)\}$ of $\mathbb{R}^N$ with $\operatorname{supp} e_j$ contained in a sphere of radius σ_T, the *windowed* WKB function $e_j(x)\varphi(x)$ is simple up to time T.

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

$$(z,s)(z',s')=\left(z+z',s+s'+\frac{[z,z']}{2}\right),$$

where $z,z'\in\mathbb{R}^{2N}$, $s,s'\in\mathbb{R}$, and

$$z=(q,p),\ z'=(q',p'),\ [z,z']=zJz'=qp'-pq',\ J=\begin{pmatrix}0 & I\\ -I & 0\end{pmatrix}.$$

For $\epsilon>0$, we define the unitary Schrödinger representation of $\mathbf{H}_N$ by

$$\tilde{\rho}_\epsilon(z,s)=\exp\left\{\frac{i}{\epsilon}(q\epsilon D-pX-sI)\right\}=\exp\left\{\frac{i}{\epsilon}([z,W]-sI)\right\},\ \rho_\epsilon(z)=\tilde{\rho}_\epsilon(z,0),$$

where $W=(X,\epsilon D)$, and D and X are two operators on distributions defined by

$$[Df](x)=-i\nabla f(x),\ [Xf](x)=xf(x). \tag{2.1}$$

For any $f\in L^2$, we have (consult page 21 in [8])

$$[\tilde{\rho}_\epsilon(z,s)f](x)=\exp\left(\frac{i}{\epsilon}\left(-px-\frac{pq}{2}-s\right)\right)f(x+q)=\exp\left(-\frac{is}{\epsilon}\right)[\rho_\epsilon(z)f](x).$$

This implies that, geometrically, $\rho_\epsilon(z)$ shifts the origin of phase space $\mathbb{R}^{2N}$ to $z=(q,p)$. If $z=z(t)$ depicts a smooth curve in the phase space, then a direct computation shows that

$$\frac{d}{dt}\rho_\epsilon(z)=\frac{i}{\epsilon}\left([\dot{z},W]-\frac{[z,\dot{z}]}{2}\right)\rho_\epsilon(z). \tag{2.2}$$

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 is defined by

$$H(W)=H(X,\epsilon D)=(2\pi\epsilon)^{-2N}\int\hat{H}_\epsilon(z)e^{\frac{i}{\epsilon}[z,W]}dz=(2\pi\epsilon)^{-2N}\int\hat{H}_\epsilon(z)\rho_\epsilon(z)dz,$$

where $\hat{H}_\epsilon$ is the ϵ-scaled symplectic Fourier transform of H defined by

$$\hat{H}_\epsilon(z)=\int H(\omega)\exp\left(\frac{i}{\epsilon}[\omega,z]\right)d\omega.$$

For any phase point z', by considering

$$\rho_\epsilon(z')\rho_\epsilon(z)\rho_\epsilon(-z')=\tilde{\rho}_\epsilon(z,-[z,z'])=e^{\frac{i}{\epsilon}[z,z']}\rho_\epsilon(z),$$

it is easy to verify that

$$\rho_\epsilon(z')H(W)\rho_\epsilon(-z')=(2\pi\epsilon)^{-2N}\int\hat{H}_\epsilon(q,p)e^{\frac{i}{\epsilon}[z,z']}\rho_\epsilon(z)dz=H(W+z'). \tag{2.3}$$

2.3. Unitary representation of unitary group. For any $f\in L^2(\mathbb{R}^N)$, its ϵ-scaled Bargmann transform is defined by

$$[B_\epsilon f](\varsigma)=\int \exp\left(-\frac{2x^2-4\varsigma x+\varsigma^2}{4\epsilon}\right) f(x)dx.$$

The Bargmann transform B_ϵ is an isometry from $L^2(\mathbb{R}^N)$ into the Fock space

$$\mathcal{F}_N=\Big\{F \;\Big|\; F \text{ is entire on } \mathbb{C}^N \\ \text{and } \|F\|_{\mathcal{F}}^2=2^{\frac{N}{2}}(2\pi\epsilon)^{-\frac{3N}{2}}\int |F(\varsigma)|^2\exp\left(-\frac{|\varsigma|^2}{2\epsilon}\right)d\varsigma<\infty\Big\}.$$

The inverse Bargmann transform from $\mathcal{F}_N$ into $L^2(\mathbb{R}^N)$ is

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

For any $U\in\mathcal{U}(N)$, the unitary transformation group in $\mathbb{C}^N$, let us define

$$\mu_\epsilon(U)=B_\epsilon^{-1}\mathcal{T}_U B_\epsilon, \qquad [\mathcal{T}_U F](\varsigma)=F(U^\dagger\varsigma),\ \forall F\in\mathcal{F}_N.$$

Since $\mathcal{T}$ gives 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 $L^2(\mathbb{R}^N)$. The following lemma reveals that the coherent state function

$$\phi(x)=\exp\left(-\frac{x^2}{2\epsilon}\right) \tag{2.4}$$

is a common fixed point of $\mu_\epsilon(U)$.

LEMMA 2.2. *For any unitary matrix U, it holds that $\mu_\epsilon(U)\phi=\phi$, where ϕ is defined by* (2.4).

Proof. A direct computation shows that

$$[B_\epsilon\phi](\varsigma)=\int \exp\left(-\frac{2x^2-4\varsigma x+\varsigma^2}{4\epsilon}\right)\exp\left(-\frac{x^2}{2\epsilon}\right)dx=(\pi\epsilon)^{\frac{N}{2}}.$$

Thus,

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

This ends the proof. ∎

Restricted to some subgroups of $\mathcal{U}(N)$, the unitary representation $\mu_\epsilon(U)$ can be made more explicit.

LEMMA 2.3. *For any real orthogonal matrix $Q\in\mathbb{R}^{N\times N}$ and any $f\in L^2(\mathbb{R}^N)$, it holds that $[\mu_\epsilon(Q^*)f](x)=f(Qx)$.*

Proof. By the definition of μ_ϵ we have

$$2^{-\frac{N}{2}}(2\pi\epsilon)^{\frac{3N}{2}}[\mu_\epsilon(Q^*)f](x) \\ =\int dy f(y)\int d\varsigma \exp\left(-\frac{2x^2-4\bar{\varsigma}x+\bar{\varsigma}^2+2|\varsigma|^2+2y^2-4(Q\varsigma)y+(Q\varsigma)^2}{4\epsilon}\right)$$

$$
\begin{aligned}
&= \int dy f(Qy) \int d\varsigma \exp\left(-\frac{2x^2-4\bar{\varsigma}x+\bar{\varsigma}^2+2|\varsigma|^2+2y^2-4\varsigma y+\varsigma^2}{4\epsilon}\right)\\
&= 2^{-\frac{N}{2}}(2\pi\epsilon)^{\frac{3N}{2}} f(Qx).
\end{aligned}
$$

This ends the proof. ∎

Lemma 2.4. *Suppose $\alpha \in \mathbb{R}^{N\times N}$ is diagonal. For any $f \in L^2(\mathbb{R}^N)$, it holds that*

$$
[\mu_\epsilon(e^{-i\alpha})f](x) = \int K_\epsilon(\alpha;x,y) f(y)dy,
$$

where

$$
K_\epsilon(\alpha;x,y) = (2\pi\epsilon)^{-\frac{N}{2}} \det\sqrt{I - i\cot\alpha}\exp\left\{\frac{i}{2\epsilon}(x\cdot\cot\alpha\cdot x - 2x\cdot\csc\alpha\cdot y + y\cdot\cot\alpha\cdot y)\right\}.
$$

Proof. By the definition of μ_ϵ we have

$$
\begin{aligned}
&2^{-\frac{N}{2}}(2\pi\epsilon)^{\frac{3N}{2}}[\mu_\epsilon(e^{-i\alpha})f](x)\\
&= \int dy f(y) \int d\varsigma \exp\left(-\frac{2x^2-4\bar{\varsigma}x+\bar{\varsigma}^2+2|\varsigma|^2+2y^2-4(e^{-i\alpha}\varsigma)y+(e^{-i\alpha}\varsigma)^2}{4\epsilon}\right)\\
&= \int dy f(y) \int d\varsigma \exp\left(-\frac{2x^2-4\bar{\varsigma}e^{-\frac{i\alpha}{2}}x+\bar{\varsigma}e^{-i\alpha}\bar{\varsigma}+2|\varsigma|^2+2y^2-4\varsigma e^{-\frac{i\alpha}{2}}y+\varsigma e^{-i\alpha}\varsigma}{4\epsilon}\right).
\end{aligned}
$$

The last equality is derived by making the change of variable $\varsigma \to e^{\frac{i\alpha}{2}}\varsigma$. Let $\varsigma = r + is$. A direct computation shows that

$$
\begin{aligned}
&\int d\varsigma \exp\left(-\frac{2x^2-4\bar{\varsigma}e^{-\frac{i\alpha}{2}}x+\bar{\varsigma}e^{-i\alpha}\bar{\varsigma}+2|\varsigma|^2+2y^2-4\varsigma e^{-\frac{i\alpha}{2}}y+\varsigma e^{-i\alpha}\varsigma}{4\epsilon}\right)\\
&= \int d\varsigma \exp\left(-\frac{x^2+y^2+r(I+e^{-i\alpha})r+s(I-e^{-i\alpha})s-2re^{-\frac{i\alpha}{2}}(x+y)+2ise^{-\frac{i\alpha}{2}}(x-y)}{2\epsilon}\right)\\
&= 2^{-\frac{N}{2}}(2\pi\epsilon)^N \det\sqrt{I-i\cot\alpha}\exp\left\{\frac{i}{2\epsilon}(x\cdot\cot\alpha\cdot x - 2x\cdot\csc\alpha\cdot y + y\cdot\cot\alpha\cdot y)\right\}.
\end{aligned}
$$

This ends the proof. ∎

Lemma 2.4 implies that $\mu_\epsilon(e^{-i\alpha})$ is actually the composition of a series of fractional Fourier transformations [5] on each pair of conjugate coordinates. For any $U \in \mathcal{U}(N)$, there exist two orthogonal matrices $Q_1, Q_2 \in \mathbb{R}^{N\times N}$ and a real diagonal matrix α such that $U = Q_1^* e^{i\alpha} Q_2$; see Theorem 2.1 and Section 5 in [30]. Given a function $f(x) \in L^2(\mathbb{R}^N)$, by Lemma 2.3 we have

$$
[\mu_\epsilon(U^*)f](x) = [\mu_\epsilon(Q_2^* e^{-i\alpha} Q_1)f](x) = [\mu_\epsilon(e^{-i\alpha})(f\circ Q_1^*)](Q_2 x).
$$

This means that, geometrically, $\mu_\epsilon(U^*)$ performs an orthogonal and symplectic transform in the phase space. In particular, it maps the Lagrangian plane U to the coordinate x-plane.

Definition 2.5. *Given a function $f(x) \in L^2(\mathbb{R}^N)$, a phase point $z = (q,p) \in \mathbb{R}^{2N}$, and a unitary matrix $U \in \mathcal{U}(N)$, the (z,U) representation of f is defined as the function*

$$
\nu_\epsilon(z,U)f = \mu_\epsilon(U^*)\rho_\epsilon(z)f.
$$

We call (z,U) a frame centered at z and directed along the Lagrangian plane U.

LEMMA 2.6. *Suppose $U=U(t)$ depicts a smooth curve in $\mathcal{U}(N)$. Set $T=\dot{U}U^*=T_R+iT_I$, where T_R and T_I are the real and imaginary parts of T respectively. Then*

$$\frac{d}{dt}\mu_\epsilon(U)=\left\{\frac{i}{2\epsilon}(XT_IX+\epsilon^2 DT_ID-\epsilon\mathrm{tr}T_I)+iXT_RD\right\}\mu_\epsilon(U).$$

Proof. It is straightforward to verify that

$$B_\epsilon^{-1}ZB_\epsilon=X-i\epsilon D,\ B_\epsilon^{-1}\nabla_\varsigma B_\epsilon=\frac{X+i\epsilon D}{2\epsilon},$$

where X and D are defined as in (2.1) by

$$[Df](x)=-i\nabla f(x),\ [Xf](x)=xf(x),$$

and

$$[ZF](\varsigma)=\varsigma F(\varsigma),\ [\nabla_\varsigma F](\varsigma)=\nabla_\varsigma F(\varsigma),\ \forall F\in\mathcal{F}_N.$$

Since

$$\frac{d}{dt}[\mathcal{T}_UF](\varsigma)=\varsigma T\nabla_\varsigma[\mathcal{T}_UF](\varsigma),$$

we have

$$\begin{aligned}\frac{d}{dt}\mu_\epsilon(U)&=B_\epsilon^{-1}(ZT\nabla_\varsigma)B_\epsilon\mu_\epsilon(U)=\frac{(X-i\epsilon D)T(X+i\epsilon D)}{2\epsilon}\mu_\epsilon(U)\\&=\frac{(X-i\epsilon D)(T_R+iT_I)(X+i\epsilon D)}{2\epsilon}\mu_\epsilon(U)\\&=\left\{\frac{i}{2\epsilon}(XT_IX+\epsilon^2DT_ID-\epsilon\mathrm{tr}T_I)+iXT_RD\right\}\mu_\epsilon(U).\end{aligned}$$

The last equality holds since T_R is anti-symmetric and T_I is symmetric. ∎

LEMMA 2.7. *For any $U\in\mathcal{U}(N)$ and $z\in\mathbb{R}^{2N}$, it holds that*

$$\mu_\epsilon(U)\rho_\epsilon(z)\mu_\epsilon(U^*)=\rho_\epsilon(R_Uz), \tag{2.5}$$

where $U=U_R+iU_I$ and

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

Proof. For any $w=r+is\in\mathbb{C}^N$, we define

$$\beta_\epsilon(w)=B_\epsilon\rho_\epsilon((r,s))B_\epsilon^{-1}.$$

For any $f\in L^2(\mathbb{R}^N)$, it holds that

$$\begin{aligned}&[\beta_\epsilon(w)B_\epsilon f](\varsigma)=[B_\epsilon\rho_\epsilon((r,s))f](\varsigma)\\&=\int dx\exp\left(-\frac{2x^2-4\varsigma x+\varsigma^2}{4\epsilon}\right)\exp\left\{-\frac{i}{\epsilon}\left(sx+\frac{sr}{2}\right)\right\}f(x+r)\\&=\int dx\exp\left(-\frac{2(x-r)^2-4\varsigma(x-r)+\varsigma^2}{4\epsilon}\right)\exp\left\{-\frac{i}{\epsilon}\left(sx-\frac{sr}{2}\right)\right\}f(x)\end{aligned}$$

$$
\begin{aligned}
&= \int dx \exp\left(-\frac{2x^2-4x(\varsigma+r-is)+2r^2+4\varsigma r+\varsigma^2-2isr}{4\epsilon}\right) f(x)\\
&= \exp\left(-\frac{|w|^2+2\varsigma w}{4\epsilon}\right)\int dx \exp\left(-\frac{2x^2-4x(\varsigma+\bar{w})+(\varsigma+\bar{w})^2}{4\epsilon}\right) f(x)\\
&= \exp\left(-\frac{|w|^2+2\varsigma w}{4\epsilon}\right)[B_\epsilon f](\varsigma+\bar{w}).
\end{aligned}
$$

This implies that for any $F \in \mathcal{F}_N$, we have

$$
[\beta_\epsilon(w)F](\varsigma) = \exp\left(-\frac{|w|^2+2\varsigma w}{4\epsilon}\right) F(\varsigma+\bar{w}).
$$

Since

$$
[\beta_\epsilon(w)T_{U^*}F](\varsigma) = \exp\left(-\frac{|w|^2+2\varsigma w}{4\epsilon}\right) F(\bar{U}(\varsigma+\bar{w})),
$$

we have

$$
\begin{aligned}
[T_U\beta_\epsilon(w)T_{U^*}F](\varsigma) &= \exp\left(-\frac{|w|^2+2(U^\dagger\varsigma)w}{4\epsilon}\right) F(\bar{U}(U^\dagger\varsigma+\bar{w}))\\
&= \exp\left(-\frac{|w|^2+2\varsigma Uw}{4\epsilon}\right) F(\varsigma+\bar{U}\bar{w}) = [\beta_\epsilon(Uw)F](\varsigma),
\end{aligned}
$$

which leads to

$$
T_U\beta_\epsilon(w)T_{U^*} = \beta_\epsilon(Uw).
$$

Let $z=(q,p)$. We then have

$$
\begin{aligned}
\mu_\epsilon(U)\rho_\epsilon(z)\mu_\epsilon(U^*) &= B_\epsilon^{-1}T_U\beta_\epsilon(q+ip)T_{U^*}B_\epsilon = B_\epsilon^{-1}\beta_\epsilon(U(q+ip))B_\epsilon\\
&= B_\epsilon^{-1}\beta_\epsilon(U_Rq-U_Ip+i(U_Iq+U_Rp))B_\epsilon\\
&= \rho_\epsilon((U_Rq-U_Ip,U_Iq+U_Rp)).
\end{aligned}
$$

This ends the proof. ∎

LEMMA 2.8. *For any $H=H(\omega)=H(x,\xi)\in\mathcal{S}'(\mathbb{R}_x^N\times\mathbb{R}_\xi^N)$, it holds that*

$$
\mu_\epsilon(U)H(W)\mu_\epsilon(U^*) = H(R_{U^*}W).
$$

Proof. Applying Lemma 2.7 we have

$$
\begin{aligned}
\mu_\epsilon(U)H(W)\mu_\epsilon(U^*) &= (2\pi\epsilon)^{-2N}\int \hat{H}_\epsilon(z)\mu_\epsilon(U)\rho_\epsilon(z)\mu_\epsilon(U^*)dz\\
&= (2\pi\epsilon)^{-2N}\int \hat{H}_\epsilon(z)\rho_\epsilon(R_Uz)dz\\
&= (2\pi\epsilon)^{-2N}\int \hat{H}_\epsilon(R_{U^*}z)\rho_\epsilon(z)dz = H(R_{U^*}W).
\end{aligned}
$$

The proof thus ends. ∎

3. Representation of simple WKB functions

Given a WKB function $\varphi(x)=A(x)\exp\left(\frac{iS(x)}{\epsilon}\right)$, let Λ be the associated Lagrangian manifold. By tangent frame we mean a frame centered at some phase point $z\in\Lambda$ and directed along the tangent plane at z.

THEOREM 3.1. *Assume that we are given a simple WKB function*

$$\varphi(x)=A(x)\exp\left(\frac{iS(x)}{\epsilon}\right),$$

where $A(x)\in\mathcal{C}_0^\infty(\mathbb{R}^N)$ and $S(x)\in\mathcal{C}^\infty(\mathbb{R}^N)$ is real. Denote by Λ the Lagrangian manifold associated with φ. For any $z=(q,p)\in\Lambda$, let U be any unitary matrix representation of the tangent plane at z. Then there exist an amplitude function $\tilde{A}(x;z)$ and a real phase function $\tilde{S}(x;z)$ on the (z,U)-plane such that, pointwise,

$$[\nu_\epsilon(z,U)\varphi](x)=[\tilde{A}(x;z)+\mathcal{O}(\epsilon)]\exp\left(\frac{i\tilde{S}(x;z)}{\epsilon}\right). \tag{3.1}$$

Moreover, $\tilde{S}(\cdot;z)$ satisfies the vanishing-derivatives property (VDP), i.e.,

$$\nabla_x\tilde{S}(0;z)=0,\ \nabla_x^2\tilde{S}(0;z)=0.$$

Furthermore, it holds that

$$\tilde{A}(0;z)=\frac{A(q)}{\det\sqrt{I-i\nabla^2S(q)}},\ \tilde{S}(0;z)=S(q)-\frac{pq}{2}.$$

Proof. The tangent plane at z is $\begin{pmatrix} I \\ \nabla^2S(q)\end{pmatrix}$. Thus $U_S=[I+(\nabla^2S(q))^2]^{-\frac{1}{2}}(I+i\nabla^2S(q))$ is a unitary matrix representation of this tangent plane. Since $\nabla^2S(q)$ is real and symmetric, there exist a real orthogonal matrix Q and a diagonal matrix α, with each entry valued in $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$, such that

$$\nabla^2S(q)=Q\tan\alpha Q^*,\ U_S=Qe^{i\alpha}Q^*.$$

For any $g\in L^2(\mathbb{R}^N)$, by Lemma 2.3 and Lemma 2.4 we have

$$[\mu_\epsilon(U_S^*)g](Qx)=[\mu_\epsilon(Q^*)\mu_\epsilon(U_S^*)g](x)=[\mu_\epsilon(e^{-i\alpha})\mu_\epsilon(Q^*)g](x)=\int K_\epsilon(\alpha;x,y)g(Qy)dy.$$

Letting

$$g(x)=[\rho_\epsilon(z)\varphi](x)=A(x+q)\exp\left(\frac{i}{\epsilon}\left[S(x+q)-p\left(x+\frac{q}{2}\right)\right]\right),$$

it holds that

$$[\mu_\epsilon(U_S^*)\rho_\epsilon(z)\varphi](Qx)=\int K_\epsilon(\alpha;x,y)A(Qy+q)\exp\left(\frac{i}{\epsilon}\left[S(Qy+q)-p\left(Qy+\frac{q}{2}\right)\right]\right)dy.$$

The above expression involves an oscillatory integral. Recalling Lemma 2.4, setting

$$T(x,y)=S(Qy+q)-p\left(Qy+\frac{q}{2}\right)+\frac{x\cdot\cot\alpha\cdot x-2x\cdot\csc\alpha\cdot y+y\cdot\cot\alpha\cdot y}{2},$$

and $V(y)=S(Qy+q)$, we have

$$\nabla_yT(x,y)=\nabla V(y)-Q^*p-(\csc\alpha)x+(\cot\alpha)y,$$

$$\nabla_y^2 T(x,y) = \nabla^2 V(y) + \cot\alpha.$$

Note that $\nabla V(0) = Q^* p$ and $\nabla^2 V(0) = \tan\alpha$. Due to the assumption of simpleness, for any x with $Qx \in \Pi_{z,U_S}\Lambda$ (the projection on the (z,U_S) plane), there exists a unique non-degenerate stationary point y_x. By applying the stationary phase theorem [28], there exist a smooth amplitude function $\hat{A}(x;z)$ and a smooth real phase function $\hat{S}(x;z)$ such that

$$[\nu_\epsilon(z,U_S)\varphi](Qx) = \left[\hat{A}(x;z) + \mathcal{O}(\epsilon)\right] \exp\left(\frac{i\hat{S}(x;z)}{\epsilon}\right). \tag{3.2}$$

Furthermore, since

$$\nabla V(y) = Q^* p + \tan\alpha \cdot y + \frac{1}{2}\nabla^3 V(0) : y^{\otimes 2} + \mathcal{O}(|y|^3),$$

when x is small it holds that

$$y_x = \cos\alpha \cdot x - \frac{\sin(2\alpha)(\nabla^3 V(0) : (\cos\alpha \cdot x)^{\otimes 2})}{4} + \mathcal{O}(|x|^3).$$

Taylor expansion shows that

$$\hat{S}(x;z) = T(x,y_x) = S(q) - \frac{pq}{2} + \mathcal{O}(|x|^3). \tag{3.3}$$

Since

$$\nabla_y^2 T(0,y_0) = \nabla^2 V(0) + \cot\alpha = \tan\alpha + \cot\alpha,$$

we have

$$\hat{A}(0;z) = A(q) \det\sqrt{I - i\cot\alpha} \cdot \det\sqrt{i\sin\alpha\cos\alpha} = \frac{A(q)}{\det\sqrt{I - i\nabla^2 S(q)}}. \tag{3.4}$$

For any unitary matrix representation U of the tangent plane at z, there exists a real orthogonal matrix $\tilde{Q}$ such that $U = U_S\tilde{Q}$. Put

$$\tilde{A}(x;z) = \hat{A}(Q^*\tilde{Q}x;z), \quad \tilde{S}(x;z) = \hat{S}(Q^*\tilde{Q}x;z).$$

By Lemma 2.3 and equation (3.2), we have

$$\begin{aligned}
[\nu_\epsilon(z,U)\varphi](x) &= [\mu_\epsilon(\tilde{Q}^*)\nu_\epsilon(z,U_S)\varphi](x) \\
&= \left[\hat{A}(Q^*\tilde{Q}x;z) + \mathcal{O}(\epsilon)\right] \exp\left(\frac{i\hat{S}(Q^*\tilde{Q}x;z)}{\epsilon}\right) \\
&= \left[\tilde{A}(x;z) + \mathcal{O}(\epsilon)\right] \exp\left(\frac{i\tilde{S}(x;z)}{\epsilon}\right).
\end{aligned}$$

According to (3.3) and (3.4), $\tilde{A}$ and $\tilde{S}$ satisfy the properties specified in the theorem. □

Remark 3.1. The geometrical meaning of VDP is obvious: the Lagrangian manifold associated with a WKB function is tangent to the coordinate x-plane at the origin of phase space. Theorem 3.1 reveals that the tangent frame representation of a simple WKB function φ satisfies three properties which are *independent* of the specific unitary matrix representation U:

- it is in the WKB form;
- the vanishing-derivatives property holds;
- the amplitude and the phase at $x=0$ are locally determined.

Note that evaluating (3.1) at $x=0$ gives a function of WKB-type on Λ:

$$[\nu_\epsilon(z,U)\varphi](0)=\left[\frac{A(q)}{\det\sqrt{I-i\nabla^2S(q)}}+\mathcal{O}(\epsilon)\right]\exp\left\{\frac{i}{\epsilon}\left(S(q)-\frac{pq}{2}\right)\right\}. \tag{3.5}$$

Its phase function is a generating function of differential 1-form $pdq-\frac{d(pq)}{2}$.

THEOREM 3.2. *Assume that we are given a WKB function (not necessarily simple)*

$$\varphi(x)=A(x)\exp\left(\frac{iS(x)}{\epsilon}\right), \tag{3.6}$$

where $A\in\mathcal{C}_0^\infty(\mathbb{R}^n)$ and $S\in\mathcal{C}^\infty(\mathbb{R}^n)$. Then the following holds pointwise:

$$\varphi(x)-(2\pi\epsilon)^{-\frac{N}{2}}\int_{z\in\Lambda}\frac{A(q)}{\det\sqrt{I-i\nabla^2S(q)}}\exp\left\{\frac{i}{\epsilon}[S(q)+\nabla S(q)(x-q)]\right\}\phi(x-q)d\mathrm{vol}=\mathcal{O}(\epsilon). \tag{3.7}$$

Here $\phi(x)$ is the coherent state function defined by (2.4).

Proof. It is straightforward to check that

$$|\exp(i\theta)-1|\le|\theta|,\ |\exp(i\theta)-1-i\theta|\le\frac{\theta^2}{\sqrt{3}},\ \forall\theta\in\mathbb{R}. \tag{3.8}$$

Besides, for any integer $m\ge 2$, there exists a universal constant $c_{m,N}>0$ such that

$$(2\pi\epsilon)^{-\frac{N}{2}}\int_{\mathbb{R}^N}|x|^m\phi(x)dx\le c_{m,N}\epsilon^{\frac{m}{2}}. \tag{3.9}$$

Setting

$$B(q)=A(q)\det\sqrt{I+i\nabla^2S(q)},$$
$$R(x,q)=S(q)+\nabla S(q)(x-q)+\frac{1}{2}\nabla^2S(x):(x-q)^{\otimes 2}-\frac{1}{3}\nabla^3S(x)\therefore(x-q)^{\otimes 3}-S(x),$$

we have

$$B(q)-(B(x)+\nabla B(x)(q-x))=\mathcal{O}(|q-x|^2),\ R(x,q)=\mathcal{O}(|q-x|^4).$$

By (3.8) we have

$$\begin{aligned}&\exp\left(\frac{i}{\epsilon}(S(q)+\nabla S(q)(x-q)-S(x))\right)\\ &=\exp\left\{\frac{i}{\epsilon}\left(-\frac{1}{2}\nabla^2S(x):(x-q)^{\otimes 2}+\frac{1}{3}\nabla^3S(x)\therefore(x-q)^{\otimes 3}+R(x,q)\right)\right\}\\ &=\exp\left(-\frac{i}{2\epsilon}\nabla^2S(x):(x-q)^{\otimes 2}\right)\left[1+\frac{i}{3\epsilon}\nabla^3S(x)\therefore(x-q)^{\otimes 3}\right]+\mathcal{O}\left(\frac{|q-x|^4}{\epsilon}+\frac{|q-x|^6}{\epsilon^2}\right).\end{aligned}$$

Furthermore, it holds that

$$B(q)\exp\left(\frac{i}{\epsilon}(S(q)+\nabla S(q)(x-q)-S(x))\right)$$

$$= [B(x)+\nabla B(x)(q-x)]\exp\left(-\frac{i}{2\epsilon}\nabla^2 S(x):(x-q)^{\otimes 2}\right)\left[1+\frac{i}{3\epsilon}\nabla^3 S(x)\therefore(x-q)^{\otimes 3}\right]$$
$$+\mathcal{O}\left(|q-x|^2+\frac{|q-x|^4}{\epsilon}+\frac{|q-x|^6}{\epsilon^2}\right)$$
$$= \left[B(x)+\nabla B(x)(q-x)+\frac{iB(x)}{3\epsilon}\nabla^3 S(x)\therefore(x-q)^{\otimes 3}\right]\exp\left(-\frac{i}{2\epsilon}\nabla^2 S(x):(x-q)^{\otimes 2}\right)$$
$$+\mathcal{O}\left(|q-x|^2+\frac{|q-x|^4}{\epsilon}+\frac{|q-x|^6}{\epsilon^2}\right).$$

By (3.9) this implies that

$$(2\pi\epsilon)^{-\frac{N}{2}}\int_{z\in\Lambda}\frac{A(q)}{\det\sqrt{I-i\nabla^2 S(q)}}\exp\left\{\frac{i}{\epsilon}[S(q)+\nabla S(q)(x-q)]\right\}\phi(x-q)d\mathrm{vol}$$
$$=(2\pi\epsilon)^{-\frac{N}{2}}\int_{\mathbb{R}^n}B(q)\exp\left\{\frac{i}{\epsilon}[S(q)+\nabla S(q)(x-q)]\right\}\phi(x-q)dq$$
$$=(2\pi\epsilon)^{-\frac{N}{2}}\exp\left(\frac{iS(x)}{\epsilon}\right)\int_{\mathbb{R}^n}B(x)\exp\left(-\frac{i}{2\epsilon}\nabla^2 S(x):(x-q)^{\otimes 2}\right)\phi(x-q)dq+\mathcal{O}(\epsilon)$$
$$=(2\pi\epsilon)^{-\frac{N}{2}}\exp\left(\frac{iS(x)}{\epsilon}\right)B(x)\frac{(2\pi\epsilon)^{\frac{N}{2}}}{\det\sqrt{I+i\nabla^2 S(x)}}+\mathcal{O}(\epsilon)$$
$$=A(x)\exp\left(\frac{iS(x)}{\epsilon}\right)+\mathcal{O}(\epsilon).$$

The proof thus finishes. □

Given a *simple* WKB function φ, according to (3.5), it holds that

$$[\nu_\epsilon(z,U)\varphi](0)=\frac{A(q)}{\det\sqrt{I-i\nabla^2 S(q)}}\exp\left(\frac{i}{\epsilon}\left[S(q)-\frac{pq}{2}\right]\right)+\mathcal{O}(\epsilon).$$

Thus,

$$\frac{A(q)}{\det\sqrt{I-i\nabla^2 S(q)}}\exp\left(\frac{i}{\epsilon}(S(q)+p(x-q))\right)\phi(x-q)$$
$$=\frac{A(q)}{\det\sqrt{I-i\nabla^2 S(q)}}\exp\left(\frac{i}{\epsilon}\left[S(q)-\frac{pq}{2}\right]\right)\exp\left(\frac{ip}{\epsilon}\left[x-\frac{q}{2}\right]\right)\phi(x-q)$$
$$=([\nu_\epsilon(z,U)\varphi](0)+\mathcal{O}(\epsilon))\,[\rho_\epsilon(-z)\phi](x).$$

By Theorem 3.2 we have

$$\varphi(x)-(2\pi\epsilon)^{-\frac{N}{2}}\int_{z\in\Lambda}[\nu_\epsilon(z,U)\varphi](0)[\rho_\epsilon(-z)\phi](x)d\mathrm{vol}=\mathcal{O}(\epsilon). \tag{3.10}$$

Since $\rho_\epsilon(-z)\phi$ expresses a coherent state with momentum $p=\nabla S(q)$ and centered at the spatial point q, the expression (3.10) reveals that with accuracy to $\mathcal{O}(\epsilon)$, one can decode a simple WKB function φ from the WKB-type function $[\nu_\epsilon(z,U)\varphi](0)$ by a coherent state integral on Λ. The formulae (3.5) and (3.10) together imply that a simple WKB function can be represented with a WKB-type function on the associated Lagrangian manifold. Different from the WKB function in the specific position representation, a WKB-type function on the Lagrangian manifold is representation independent. Moreover, the latter can be naturally generalized to Lagrangian manifolds which do not admit a single-valued projection onto the position q-plane, as shown in the following lemma.

LEMMA 3.3. *Suppose* Λ *is an* N*-dimensional simple Lagrangian manifold, and assume that we are given a WKB-type function* $A(z)\exp\left(\frac{iS(z)}{\epsilon}\right)$ *with* $A\in\mathcal{C}_0^\infty(\Lambda)$ *and* $S\in\mathcal{C}^\infty(\Lambda)$*. Suppose* S *is a generating function of differential 1-form* $pdq-\frac{d(pq)}{2}$ *on* Λ*. Define*

$$\varphi(x)=(2\pi\epsilon)^{-\frac{N}{2}}\int_{z\in\Lambda}A(z)\exp\left(\frac{iS(z)}{\epsilon}\right)[\rho_\epsilon(-z)\phi](x)d\mathrm{vol},$$

where ϕ *is the coherent state function defined by* (2.4)*. Then*

1. *If* Λ *is diffeomorphically projected to the* q*-plane, then* φ *is a WKB function whose associated Lagrangian manifold is* Λ*;*
2. *In the general case, for any* $z_f\in\operatorname{supp}A$*, the* (z_f,U_f) *representation of* φ*, i.e., the function* $[\nu_\epsilon(z_f,U_f)\varphi](x)$*, is a WKB function of VDP whose associated Lagrangian manifold is* Λ*. In addition,*

$$[\nu_\epsilon(z_f,U_f)\varphi](0)=A(z_f)\exp\left(\frac{iS(z_f)}{\epsilon}\right).$$

Proof. If Λ is diffeomorphically projected to the q-plane, then q is a coordinate of Λ. Since S is a generating function of $pdq-\frac{d(pq)}{2}$ on Λ, the function $\tilde S(q)=S(z)+\frac{pq}{2}$ is a generating function of pdq. This means $\nabla\tilde S(q)=p$. Set

$$\tilde A(q)=A(z)\det\sqrt{I-i\nabla^2\tilde S(q)},\ \tilde\varphi(x)=\tilde A(x)\exp\left(\frac{i\tilde S(x)}{\epsilon}\right).$$

The Lagrangian manifold of $\tilde\varphi$ is Λ, and by (3.5), it holds that

$$[\nu_\epsilon(z,U)\tilde\varphi](0)=A(z)\exp\left(\frac{iS(z)}{\epsilon}\right).$$

Applying (3.10) gives

$$\tilde\varphi(x)-\varphi(x)=\mathcal{O}(\epsilon),$$

which means that φ is a WKB function with Lagrangian manifold Λ. The first assertion is proved.

In the general case, let us set

$$Z=(Q,P)=Z(z)=R_{U_f^*}(z-z_f).$$

A direct computation shows that

$$[\nu_\epsilon(z_f,U_f)\varphi](x)=(2\pi\epsilon)^{-\frac{N}{2}}\int_{z\in\Lambda}A(z)\exp\left(\frac{iS(z)}{\epsilon}\right)\exp\left(\frac{i[z_f,z]}{2\epsilon}\right)[\rho_\epsilon(-Z)\phi](x)d\mathrm{vol}.$$

Since $z=(q,p)\to Z=(Q,P)$ is a canonical transformation, $Z(\Lambda)$ is Lagrangian in the Z coordinate. It is straightforward to verify that

$$pdq-\frac{d(pq)}{2}+\frac{d[z_f,z]}{2}=PdQ-\frac{d(PQ)}{2}.$$

Since S is a generating function of $pdq - \frac{d(pq)}{2}$ on Λ, $S(z) + \frac{[z_f,z]}{2}$ is a generating function of $PdQ - \frac{d(PQ)}{2}$ on $Z(\Lambda)$ in the Z coordinate. By the first assertion, $\nu_\epsilon(z_f,U_f)\varphi$ is a WKB function whose Lagrangian manifold is $Z(\Lambda)$. Since $Z(\Lambda)$ is tangent to the Q-plane at $Z=0$, i.e., $z=z_f$, we have

$$[\nu_\epsilon(z_f,U_f)\varphi](0) = A(z_f)\exp\left(\frac{iS(z_f)}{\epsilon}\right).$$

This ends the proof. □

4. Global asymptotic solution for the Schrödinger equation

In this section, we consider the Schrödinger equation of the form

$$i\epsilon\partial_t u = H(X,\epsilon D,t)u, \tag{4.1}$$

$$u(x,0) = u_I(x) = A_I(x)\exp\left(\frac{iS_I(x)}{\epsilon}\right). \tag{4.2}$$

First let us assume that the initial WKB data (4.2) is simple up to time $T>0$.

4.1. WKB analysis in the moving frame. Given an initial phase point z_I, the flow induced by the Hamiltonian $H(z,t)$ renders a path $z(t)$ which obeys the Hamiltonian system

$$\dot{z} = J\nabla H(z,t),\ z(0) = z_I.$$

Besides, given a Lagrangian N-plane $C_I \in \mathbb{R}^{2N\times N}$, the flow also renders a continuous set of Lagrangian N-planes $C(t)$ which solves

$$\dot{C} = J\nabla^2 H(z,t)C,\ C(0) = C_I.$$

Let $C=QP$ be the polar decomposition. Then the orthogonal matrix Q represents the same Lagrangian N-plane as C does. Partitioning Q as $\begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix}$, we know that $U = Q_1 + iQ_2 \in \mathcal{U}(N)$ is a unitary matrix representation of C. We call (z,U) a *moving frame* associated with the Hamiltonian H. It should be pointed out that the time dependence has been made implicit, mainly for the sake of notation. Now we attempt to express the Schrödinger equation (4.1) in this moving frame. More precisely, by setting

$$f(x,t) = [\nu_\epsilon(z,U)u](x) = [\mu_\epsilon(U^*)\rho_\epsilon(z)u](x),$$

we want to derive an equation satisfied by f. Since $u = \rho_\epsilon(-z)\mu_\epsilon(U)f$, a direct computation shows that

$$\begin{aligned} i\epsilon\partial_t f &= i\epsilon\left(\frac{d}{dt}\mu_\epsilon(U^*)\right)\rho_\epsilon(z)u + i\epsilon\mu_\epsilon(U^*)\left(\frac{d}{dt}\rho_\epsilon(z)\right)u + i\epsilon\mu_\epsilon(U^*)\rho_\epsilon(z)\partial_t u \\ &= i\epsilon\left(\frac{d}{dt}\mu_\epsilon(U^*)\right)\mu_\epsilon(U)f + i\epsilon\mu_\epsilon(U^*)\left(\frac{d}{dt}\rho_\epsilon(z)\right)\rho_\epsilon(-z)\mu_\epsilon(U)f \\ &\quad + \mu_\epsilon(U^*)\rho_\epsilon(z)H(W,t)\rho_\epsilon(-z)\mu_\epsilon(U)f. \end{aligned} \tag{4.3}$$

By Lemma 2.6, Lemma 2.7, and formulae (2.3) and (2.2), we have

$$i\epsilon\left(\frac{d}{dt}\mu_\epsilon(U^*)\right)\mu_\epsilon(U) = -\frac{1}{2}(XT_IX + \epsilon^2 DT_ID - \epsilon\mathrm{tr}T_I) - \epsilon XT_RD,\ T_R + iT_I = \dot{U}^*U,$$

$$i\epsilon\mu_\epsilon(U^*)\left(\frac{d}{dt}\rho_\epsilon(z)\right)\rho_\epsilon(-z)\mu_\epsilon(U)=-[\dot z,R_UW]+\frac{[z,\dot z]}{2}=-R_UW\cdot\nabla H(z,t)+\frac{[z,\dot z]}{2},$$
$$\mu_\epsilon(U^*)\rho_\epsilon(z)H(W,t)\rho_\epsilon(-z)\mu_\epsilon(U)=H(R_UW+z,t).$$

Substituting the above into (4.3) we arrive at

$$i\epsilon\partial_t f= \left[H(R_UW+z,t)-R_UW\cdot\nabla H(z,t)+\frac{[z,\dot z]}{2}-\frac{1}{2}(XT_IX+\epsilon^2DT_ID-\epsilon\mathrm{tr}T_I)-\epsilon XT_RD\right]f. \tag{4.4}$$

Let us introduce

$$\tilde H(\omega,t,\delta)=H(R_U\omega+z,t)-R_U\omega\cdot\nabla H(z,t)+\frac{[z,\dot z]}{2}-\frac{1}{2}(xT_Ix+\xi T_I\xi)-xT_R\xi+\frac{i\delta}{2}\mathrm{tr}T_I.$$

Equation (4.4) can be rewritten into an abbreviated form:

$$-i\epsilon\partial_t f+\tilde H(W,t,-i\epsilon)f=0. \tag{4.5}$$

Now we make a standard WKB analysis for (4.5). Suppose f has the asymptotic expansion

$$f=\exp\left(\frac{iS}{\epsilon}\right)\sum_{j=0}^{\infty}(-i\epsilon)^j\varphi_j. \tag{4.6}$$

A direct computation shows that

$$-i\epsilon\partial_t f=\exp\left(\frac{iS}{\epsilon}\right)\left[\sum_{j=1}^{\infty}(-i\epsilon)^j\partial_t\varphi_{j-1}+\sum_{j=0}^{\infty}(-i\epsilon)^j\varphi_j\partial_t S\right]. \tag{4.7}$$

Besides, we have

$$\tilde H(W,t,-i\epsilon)\left(\varphi\exp\left(\frac{iS}{\epsilon}\right)\right)=\exp\left(\frac{iS}{\epsilon}\right)\sum_{j=0}^{\infty}(-i\epsilon)^jR_j[\varphi],$$

where R_j are linear differential operators of order j. In particular,

$$R_0[\varphi]=\tilde H\varphi, \tag{4.8}$$
$$R_1[\varphi]=\nabla_2\tilde H\cdot\nabla\varphi+\frac{\nabla_2^2\tilde H:\nabla^2 S}{2}\varphi+\frac{\mathrm{tr}(\nabla_1\nabla_2\tilde H)}{2}\varphi+\partial_\delta\tilde H\varphi. \tag{4.9}$$

Here ∇_k denotes the derivative with respect to the k-th placeholder variable, and the function $\tilde H$ is evaluated at $(x,\nabla S,t,0)$. Thus,

$$\tilde H(W,t,-i\epsilon)f =\exp\left(\frac{iS}{\epsilon}\right)\sum_{j=0}^{\infty}(-i\epsilon)^j\left[\sum_{k=0}^{\infty}(-i\epsilon)^kR_k[\varphi_j]\right]=\exp\left(\frac{iS}{\epsilon}\right)\sum_{j=0}^{\infty}(-i\epsilon)^j\left[\sum_{k=0}^{j}R_k[\varphi_{j-k}]\right]. \tag{4.10}$$

Adding up (4.7) and (4.10) gives

$$-i\epsilon\partial_t f+\tilde H(W,t,-i\epsilon)f$$

$$
\begin{aligned}
&= \left(\sum_{j=1}^{\infty} (-i\epsilon)^j \partial_t \varphi_{j-1} + \sum_{j=0}^{\infty} (-i\epsilon)^j \varphi_j \partial_t S + \sum_{j=0}^{\infty} (-i\epsilon)^j \left[\sum_{k=0}^{j} R_k[\varphi_{j-k}] \right] \right) \exp\left(\frac{iS}{\epsilon} \right) \\
&\equiv \left(\Box_0 + (-i\epsilon)\Box_1 + (-i\epsilon)^2 \Box_2 + \cdots \right) \exp\left(\frac{iS}{\epsilon} \right),
\end{aligned}
$$

where

$$
\begin{aligned}
\Box_0 &= \varphi_0 \left(\partial_t S + \tilde{H}(x, \nabla S, t, 0) \right), \\
\Box_1 &= \partial_t \varphi_0 + R_1[\varphi_0] + \varphi_1 \left(\partial_t S + \tilde{H}(x, \nabla S, t, 0) \right), \\
\Box_j &= \partial_t \varphi_{j-1} + R_1[\varphi_{j-1}] + \varphi_j \left(\partial_t S + \tilde{H}(x, \nabla S, t, 0) \right) + \sum_{k=2}^{j} R_k[\varphi_{j-k}], \ \forall j \geq 2.
\end{aligned}
$$

If S solves the phase equation

$$
\partial_t S + \tilde{H}(x, \nabla S, t, 0) = 0,
$$

and the φ_j solve the transport equations

$$
\begin{aligned}
&\partial_t \varphi_0 + R_1[\varphi_0] = 0, \\
&\partial_t \varphi_{j-1} + R_1[\varphi_{j-1}] + \sum_{k=2}^{j} R_k[\varphi_{j-k}] = 0, \ \forall j \geq 2,
\end{aligned}
$$

we then derive a solution (4.6) to equation (4.5) in the asymptotic series form.

4.2. WKB solution in a specific moving frame. Let Λ_I be the Lagrangian manifold associated with the WKB initial data u_I; see (4.2). Suppose the moving frame initiates from a specific frame (z_I, U_I) with $z_I \in \Lambda_I$ and U_I being an equivalent unitary matrix representation for the tangent plane $C_I = \begin{pmatrix} I \\ \nabla^2 S_I(q_I) \end{pmatrix}$. By Theorem 3.1, there exist an amplitude function $A(x, 0; z_I)$ and a phase function $S(x, 0; z_I)$ of VDP, such that

$$
f(x, 0) = [\nu_\epsilon(z_I, U_I) u_I](x) = (A(x, 0; z_I) + \mathcal{O}(\epsilon)) \exp\left(\frac{iS(x, 0; z_I)}{\epsilon} \right),
$$

with

$$
A(0, 0; z_I) = \frac{A_I(q_I)}{\det \sqrt{I - i\nabla^2 S_I(q_I)}}, \ S(0, 0; z_I) = S_I(q_I) - \frac{p_I q_I}{2}.
$$

The function $f(x, 0)$ is in the WKB form, and we can apply the WKB analysis in the last subsection to seek a solution of (4.5) with the ansatz

$$
f(x, t) = (A(x, t; z_I) + \mathcal{O}(\epsilon)) \exp\left(\frac{iS(x, t; z_I)}{\epsilon} \right). \tag{4.11}
$$

Here S and A solve the following equations:

$$
\partial_t S + \tilde{H}(x, \nabla_x S, t, 0) = 0, \tag{4.12}
$$

$$
\partial_t A + R_1[A] = 0. \tag{4.13}
$$

Let g^t denote the displacement operator associated with the Hamiltonian H, and set $\Lambda(t)=g^t\Lambda_I$. Since g^t is canonical, $\Lambda(t)$ is Lagrangian. Set

$$w=w(\xi)=R_{U^*}(\xi-z),\ \forall\xi\in\mathbb{R}^{2N}.$$

Then $\xi\to w$ is a linear canonical transformation for any $t\in[0,T]$.

THEOREM 4.1. *Suppose the initial WKB data* (4.2) *is simple up to time* $T>0$*, and the moving frame* (z,U) *is determined by*

$$\begin{aligned}&\dot z=J\nabla H(z,t),\ z(0)=z_I=(q_I,p_I)\in\Lambda_I,\\ &\dot C=J\nabla^2H(z,t)C,\ C(0)=\begin{pmatrix}I\\ \nabla^2S_I(q_I)\end{pmatrix},\\ &C=QP,\ Q=\begin{pmatrix}Q_1\\ Q_2\end{pmatrix},\ U=Q_1+iQ_2.\end{aligned}$$

Then the following holds:

1. *the solution* $S(x,t;z_I)$ *of* (4.12) *exists on the domain* $\Pi_{z,U}\Lambda$ *of the* (z,U)*-plane for any* $t\in[0,T]$*;*
2. *the Lagrangian manifold associated with* $S(x,t;z_I)$ *on* (z,U)*-plane is* Λ*;*
3. *the vanishing-derivatives property holds, i.e.,*

$$\nabla_xS(0,t;z_I)=0,\ \nabla_x^2S(0,t;z_I)=0;$$

4. $S(0,t;z_I)$ *solves the ODE*

$$\partial_tS(0,t;z_I)+H(z,t)+\frac{[z,\dot z]}{2}=0,\ S(0,0;z_I)=S_I(q_I)-\frac{p_Iq_I}{2}.$$

5. $A(0,t;z_I)$ *solves the ODE*

$$\begin{aligned}\partial_tA(0,t;z_I)+\frac{\operatorname{tr}(Q^*\nabla^2H(z,t)(iI-J)Q)}{2}A(0,t;z_I)=0\\ A(0,0;z_I)=\frac{A_I(q_I)}{\det\sqrt{I-i\nabla^2S_I(q_I)}}.\end{aligned}$$

Proof. The Hamiltonian system corresponding to (4.12) is

$$\dot w=J\left(R_{U^*}\nabla H(R_Uw+z,t)-R_{U^*}\nabla H(z,t)-\begin{pmatrix}T_I & T_R\\ T_R & T_I\end{pmatrix}w\right).$$

Since

$$T_R+iT_I=\dot U^*U=-U^*\dot U,$$

a direct computation shows that

$$\begin{pmatrix}T_R & T_I\\ -T_I & T_R\end{pmatrix}=R_{U^*}\frac{dR_U}{dt}.$$

Changing to the ξ-coordinate with

$$\xi=z+R_Uw,$$

we have

$$\begin{aligned}\dot{\xi} &= \dot{z} + \frac{d(R_U w)}{dt} = J\nabla H(z,t) + R_U \dot{w} + \frac{dR_U}{dt} w \\ &= J\nabla H(\xi,t) + \left(\frac{dR_U}{dt} - R_U J \begin{pmatrix} T_I & T_R \\ T_R & T_I \end{pmatrix} \right) w \\ &= J\nabla H(\xi,t) + \left(\frac{dR_U}{dt} - R_U \begin{pmatrix} T_R & T_I \\ -T_I & T_R \end{pmatrix} \right) w = J\nabla H(\xi,t),\end{aligned}$$

which is the Hamiltonian system with the Hamiltonian H. Since $\xi(0)$ is valued on Λ_I, we know the Lagrangian manifold of w is simply $\Lambda(t)$. Thus $S(x,t;z_I)$ is single-valued on $\Pi_{z,U}\Lambda$ for any $t \in [0,T]$, since we have assumed Λ_I is simple up to T. The Lagrangian manifold Λ is tangent to the (z,U)-plane at $x=0$, which implies the third assertion. Finally, evaluating (4.12) at $x=0$ gives the ODE satisfied by $S(0,t;z_I)$.

Next let us consider equation (4.13). Evaluating it at $x=0$ and applying the vanishing-derivatives property of $S(x,t;z_I)$, we have

$$\partial_t A(0,t;z_I) + \frac{\mathrm{tr}(\nabla_1 \nabla_2 \tilde{H}(0,0,t,0))}{2} A(0,t;z_I) + \partial_\delta \tilde{H}(0,0,t,0) A(0,t;z_I) = 0.$$

A direct computation shows that

$$\nabla^2 \tilde{H}(w,t,0) = R_{U^*} \nabla^2 H(R_U w + z,t) R_U - \begin{pmatrix} T_I & T_R \\ T_R & T_I \end{pmatrix}.$$

Thus,

$$\nabla^2 \tilde{H}(0,0,t,0) = R_{U^*} \nabla^2 H(z,t) R_U - \begin{pmatrix} T_I & T_R \\ T_R & T_I \end{pmatrix}.$$

Since $R_U = (Q, -JQ)$, we have

$$\begin{aligned}\nabla^2 \tilde{H}(0,0,t,0) &= \begin{pmatrix} Q^* \\ Q^* J \end{pmatrix} \nabla^2 H(z,t)(Q,-JQ) - \begin{pmatrix} T_I & T_R \\ T_R & T_I \end{pmatrix} \\ &= \begin{pmatrix} Q^*\nabla^2 H(z,t) Q - T_I & -Q^*\nabla^2 H(z,t) JQ - T_R \\ Q^* J \nabla^2 H(z,t) Q - T_R & -Q^* J \nabla^2 H(z,t) JQ - T_I \end{pmatrix}.\end{aligned}$$

Since T_R is anti-symmetric, it holds that

$$\mathrm{tr}(\nabla_1 \nabla_2 \tilde{H}(0,0,t,0)) = -\mathrm{tr}(Q^* \nabla^2 H(z,t) JQ).$$

Considering

$$\partial_\delta \tilde{H}(0,0,t,0) = \frac{i\,\mathrm{tr} T_I}{2},$$

to prove the last assertion it suffices to show that

$$T_I = Q^* \nabla^2 H(z,t) Q. \tag{4.14}$$

Since

$$C_t = J\nabla^2 H(z,t) C$$

and $C = QP$, we have

$$J\nabla^2 H(z,t) QP = Q_t P + QP_t.$$

Thus,

$$J\nabla^2 H(z,t)Q = Q_t + QP_tP^{-1}.$$

Multiplying both sides with $-Q^*J$ gives

$$Q^*\nabla^2 H(z,t)Q = -Q^*JQ_t - Q^*JQP_tP^{-1}.$$

Considering that $Q^*JQ=0$ due to the symplectic property, we then have

$$Q^*\nabla^2 H(z,t)Q = -Q^*JQ_t. \tag{4.15}$$

Since

$$U_t^*U = (Q_{1,t}^* - iQ_{2,t}^*)(Q_1+iQ_2) = (Q_{1,t}^*Q_1 + Q_{2,t}^*Q_2) + i(Q_{1,t}^*Q_2 - Q_{2,t}^*Q_1),$$

it holds that

$$T_R = Q_{1,t}^*Q_1 + Q_{2,t}^*Q_2,\ T_I = Q_{1,t}^*Q_2 - Q_{2,t}^*Q_1.$$

Considering that T_I is symmetric, we have

$$T_I = Q_2^*Q_{1,t} - Q_1^*Q_{2,t} = -Q^*JQ_t. \tag{4.16}$$

Combining (4.15) and (4.16) yields (4.14). ∎

Remark 4.1. From the proof of Theorem 4.1, it can be seen that if f is a WKB function of VDP for any $t\in[0,T]$, then at $x=0$ it holds that

$$\left(-i\epsilon\partial_t f + \tilde{H}(W,t,-i\epsilon)f\right)_{x=0} = (\Box_0|_{x=0} + (-i\epsilon)\Box_1|_{x=0} + \cdots)\exp\left(\frac{iS(0,t)}{\epsilon}\right),$$

with

$$\begin{aligned}
\Box_0|_{x=0} &= \varphi_0\left(\partial_t S(0,t) + H(z,t) + \frac{[z,\dot{z}]}{2}\right),\\
\Box_1|_{x=0} &= \partial_t\varphi_0(0,t) + \frac{\operatorname{tr}(Q^*\nabla^2 H(z,t)(iI-J)Q)}{2}\varphi_0(0,t)\\
&\quad + \varphi_1(0,t)\left(\partial_t S(0,t) + H(z,t) + \frac{[z,\dot{z}]}{2}\right).
\end{aligned}$$

4.3. Global asymptotic solution. The last subsection has shown that if we express the Schrödinger equation in a specific moving frame started with (z_I,U_I), a WKB solution is available for any $t\in[0,T]$. Besides, the phase and the amplitude at $x=0$ can be determined by simple ODEs. Now let us change z_I and take z_I as an argument in the initial Lagrangian manifold Λ_I. For any $z=g^t z_I\in\Lambda(t)=g^t\Lambda_I$, by Theorem 4.1 we have

$$[\nu_\epsilon(z,U)u](0) = [A(t;z_I) + \mathcal{O}(\epsilon)]\exp\left(\frac{iS(t;z_I)}{\epsilon}\right), \tag{4.17}$$

which expresses a WKB function living only on the Lagrangian manifold Λ. Besides, A and S are determined by

$$\dot{S} + H(z,t) + \frac{[z,\dot{z}]}{2} = 0,\ S(0;z_I) = S_I(q_I) - \frac{p_Iq_I}{2}, \tag{4.18}$$

$$\dot{A}+\frac{\mathrm{tr}(Q^*\nabla^2 H(z,t)(iI-J)Q)}{2}A=0,\ A(0;z_I)=\frac{A_I(q_I)}{\det\sqrt{I-i\nabla^2 S_I(q_I)}}. \tag{4.19}$$

Note that for the sake of brevity, we have rewritten $A(0,t;z_I)$ into $A(t;z_I)$, and $S(0,t;z_I)$ into $S(t;z_I)$. The WKB-type function (4.17) encodes the solution u of (4.1)-(4.2) through a continuous set of representations parameterized by the Lagrangian manifold Λ. In the following, we show that u can be decoded from (4.17) with accuracy to $\mathcal{O}(\epsilon)$.

According to the WKB analysis in the last subsection, for a specific moving frame (z_f,U_f), the function $[\nu_\epsilon(z_f,U_f)u](x)$ is a WKB function on the (z_f,U_f)-plane for any $t\in[0,T]$. Besides, the associated Lagrangian manifold is simply Λ. Set

$$Z=(Q,P)=R_{U_f^*}(z-z_f).$$

Applying (3.10) and Lemma 2.7, we have (with accuracy to $\mathcal{O}(\epsilon)$)

$$\begin{aligned}[\nu_\epsilon(z_f,U_f)u](x)&=(2\pi\epsilon)^{-\frac{N}{2}}\int_{z\in\Lambda}[\nu_\epsilon(Z,U_f^*U)\nu_\epsilon(z_f,U_f)u](0)[\rho_\epsilon(-Z)\phi](x)d\mathrm{vol}\\&=(2\pi\epsilon)^{-\frac{N}{2}}\int_{z\in\Lambda}\exp\left(-\frac{i[z,z_f]}{2\epsilon}\right)[\nu_\epsilon(z,U)u](0)[\rho_\epsilon(-Z)\phi](x)d\mathrm{vol},\end{aligned}$$

where ϕ is the coherent state function defined by (2.4). Then, by Lemma 2.2 and Lemma 2.7, it holds that

$$[\rho_\epsilon(z_f)u](x)=(2\pi\epsilon)^{-\frac{N}{2}}\int_\Lambda\exp\left(-\frac{i[z,z_f]}{2\epsilon}\right)[\nu_\epsilon(z,U)u](0)[\rho_\epsilon(z_f-z)\phi](x)d\mathrm{vol}.$$

Furthermore, by the group property of ρ_ϵ we have

$$u(x,t)=(2\pi\epsilon)^{-\frac{N}{2}}\int_\Lambda[\mu_\epsilon(z,U)u](0)[\rho_\epsilon(-z)\phi](x)d\mathrm{vol}.$$

Substituting (4.17) into the above formula yields

$$\boxed{u(x,t)=(2\pi\epsilon)^{-\frac{N}{2}}\int_\Lambda A\exp\left(\frac{iS}{\epsilon}\right)[\rho_\epsilon(-z)\phi](x)d\mathrm{vol}.} \tag{4.20}$$

The formula (4.20) gives an intrinsic coherent state approximation to the wave field u with accuracy to $\mathcal{O}(\epsilon)$. Though we derive it based on a fixed moving frame (z_f,U_f), the final result does not rely on this choice. In case the initial WKB data (4.2) is not simple, we can find a partition of unity $\{e_j(x)\}$ of $\mathbb{R}^N$ such that each WKB wave function $u_I^j(x)=e_j(x)u_I(x)$ is simple up to time T. The Lagrangian manifold Λ is then decomposed into a union of sub-manifolds Λ^j. Solving the Schrödinger equation (4.1) with the initial WKB data u_I^j by the proposed method, we derive a first order approximation of the wave field u^j by replacing Λ with Λ^j in (4.20). Due to the linearity property of the Schrödinger equation, a first order approximation of the total wave field u is then derived by summing up u^j together. The final output is expressed exactly in the same form as (4.20). For ease of reference, we reformulate the ODEs satisfied by A and S, together with the Hamiltonian system and the equation for the tangent plane, into the following system:

$$\dot{z}=J\nabla H(z,t),\ z(0;z_I)=(q_I,\nabla S_I(q_I)), \tag{4.21}$$

$$\dot{C}=J\nabla^2 H(z,t)C,\ C(0;z_I)=\begin{pmatrix} I \\ \nabla^2 S_I(q_I)\end{pmatrix},\ C=QP, \tag{4.22}$$

$$\dot{S}+H(z,t)+\frac{[z,\dot{z}]}{2}=0,\ S(0;z_I)=S_I(q_I)-\frac{p_I q_I}{2}, \tag{4.23}$$

$$\dot{A}+\frac{\mathrm{tr}(Q^*\nabla^2 H(z,t)(iI-J)Q)}{2}A=0,\ A(0;z_I)=\frac{A_I(q_I)}{\det\sqrt{I-i\nabla^2 S_I(q_I)}}. \tag{4.24}$$

5. Cauchy problem of general scalar wave equation

In this section, we consider the scalar wave equation of the form

$$H(X,\epsilon D)u=0. \tag{5.1}$$

We should require that the Hamiltonian flow associated with H does not admit any finite motion. This means that for any bounded domain in $\mathbb{R}^N$, any ray curve will remain outside of this bounded domain after a finite evolution time. A typical example of the equation (5.1), apart from the Schrödinger equation considered in the last section, is the Helmholtz equation in a heterogeneous medium:

$$\Delta u+k^2 n(x)u=0,\ x\in\mathbb{R}^N.$$

In this case, we have $\epsilon=1/k$ and $H(x,\xi)=\xi^2-n(x)$.

5.1. Specifying the Cauchy data. The Cauchy data is prescribed on a smooth orientable manifold Ω of dimension $(N-1)$ in $\mathbb{R}^N$ as

$$u_I(q_I)=A_I(q_I)\exp\left(\frac{iS_I(q_I)}{\epsilon}\right),\ q_I\in\Omega. \tag{5.2}$$

Here A_I and S_I are smooth functions merely on Ω. Some requirement of uniform regularity may also be needed, which will be assumed to hold whenever necessary.

Let us first try to seek a local WKB solution of (5.1) with the Cauchy data (5.2) in a small neighborhood of Ω as

$$u(x)=A(x)\exp\left(\frac{iS(x)}{\epsilon}\right). \tag{5.3}$$

By the WKB analysis, the phase function S should satisfy the eikonal equation

$$H(q,\nabla S(q))=0.$$

Restricted to Ω, we have

$$H(q_I,\nabla S(q_I))=H(q_I,\nabla_\Omega S_I(q_I)+\nabla_\Omega^\perp S(q_I))=0. \tag{5.4}$$

Here ∇_Ω denotes the covariant derivative and $\nabla_\Omega^\perp$ is the normal derivative. Since Ω is of co-dimension 1, in principle $\nabla_\Omega^\perp S(q_I)$ can be determined by solving the algebraic equation (5.4). However, $\nabla_\Omega^\perp S(q_I)$ might be multi-valued which typically relates to different wave directions. Therefore, in general we need also to specify $p_I=\nabla S(q_I)$ such that with respect to q_I, p_I is smooth and single-valued in Ω. With this been done, the Cauchy data (5.2) renders a smooth Lagrangian manifold of dimension $N-1$ as

$$\Lambda_{N-1}^I=\{(q_I,p_I)\ |\ q_I\in\Omega\}. \tag{5.5}$$

Let g^t indicate the displacement operator induced by the Hamiltonian system

$$\dot{z} = J\nabla H(z).$$

If the velocity $\nabla_p H(q_I, p_I)$ is always transversal to the tangent plane for each $q_I \in \Omega$, the Hamiltonian flow renders an N-dimensional Lagrangian manifold

$$\Lambda = \bigcup_{t\in\mathbb{R}} g^t \Lambda^I_{N-1},$$

which is topologically isomorphic to $\Omega \times \mathbb{R}$. By Theorem 5.3 in [28], there exists a local WKB solution (5.3) which is compatible with the Cauchy initial data.

Now we determine a matrix representation C_I of the tangent plane at $z_I = (q_I, p_I) \in \Lambda^I_{N-1} \subset \Lambda$. Let $y \in \mathbb{R}^{N-1}$ be a local coordinate in a small connected open set which contains q_I. Then (y,t) is a local coordinate in a small neighborhood of z_I on Λ. Since

$$\nabla^2 S = \frac{\partial p}{\partial q} = \frac{\partial y}{\partial q}\frac{\partial p}{\partial y} + \frac{\partial t}{\partial q}\frac{\partial p}{\partial t} = \frac{\partial y}{\partial q}\frac{\partial p}{\partial y} - \frac{\partial t}{\partial q} H_q, \tag{5.6}$$

where $(\frac{\partial y}{\partial q}, \frac{\partial t}{\partial q})$ is the inverse matrix of

$$\begin{pmatrix} \frac{\partial q}{\partial y} \\ \frac{\partial q}{\partial t} \end{pmatrix} = \begin{pmatrix} \frac{\partial q}{\partial y} \\ H_p \end{pmatrix},$$

evaluating (5.6) at $q = q_I$ we have

$$\nabla^2 S(q_I) = \begin{pmatrix} \frac{\partial q_I}{\partial y} \\ H_p(q_I, p_I) \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial p_I}{\partial y} \\ -H_q(q_I, p_I) \end{pmatrix}. \tag{5.7}$$

The invertibility of the matrix in the above expression is ensured by the transversality assumption. The Lagrangian plane C_I is then set as

$$C_I = \begin{pmatrix} I \\ \nabla^2 S(q_I) \end{pmatrix}.$$

Along each trajectory, the Hamiltonian flow induces a smooth transformation of the tangent plane C as

$$\dot{C} = J\nabla^2 H(z) C.$$

Let $C = QP$ with $Q = \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix}$ being the polar decomposition. Then $U = Q_1 + iQ_2$ is a unitary matrix representation of tangent plane C.

5.2. Formal derivation of the global asymptotic solution. Suppose $z = z(t)$ is a path started with $z_I \in \Lambda_I$, and $C = C(t)$ is the Lagrangian plane started with C_I. For any suitable function $u \in \mathcal{S}'(\mathbb{R}^N_x)$, let us introduce

$$f(x,t) = [\nu_\epsilon(z,U)u](x).$$

The function f can be taken as an $\mathbb{R}$-copied version of u in the moving frame (z,U). Analogous to the derivation of (4.4), we have

$$i\epsilon\partial_t f = \left[-R_U W \cdot \nabla H(z) + \frac{[z,\dot{z}]}{2} - \frac{1}{2}(XT_I X + \epsilon^2 D T_I D - \epsilon \operatorname{tr} T_I) - \epsilon X T_R D \right] f, \tag{5.8}$$

where $T_R + iT_I = \dot{U}^* U$. If u is a solution of (5.1), namely,

$$H(W)u = H(W)\rho_\epsilon(-z)\mu_\epsilon(U)f = 0,$$

we have

$$H(R_U W + z)f = \mu_\epsilon(U^*)\rho_\epsilon(z)H(W)\rho_\epsilon(-z)\mu_\epsilon(U)f = 0. \tag{5.9}$$

Adding up (5.8) and (5.9) gives

$$\begin{aligned} & i\epsilon\partial_t f \\ &= \left[H(R_U W + z) - R_U W \cdot \nabla H(z) + \frac{[z,\dot{z}]}{2} - \frac{1}{2}(X T_I X + \epsilon^2 D T_I D - \epsilon \mathrm{tr} T_I) - \epsilon X T_R D\right] f. \end{aligned} \tag{5.10}$$

Apart from the explicit time dependence getting involved in the Hamiltonian function, the equation (5.10) is the same as (4.4).

Suppose the Lagrangian manifold Λ is simple and u is a global WKB solution compatible with the Cauchy initial data. Then for any $t \in \mathbb{R}$, $f(x,t)$ is a WKB function by Theorem 3.1. Of course, these WKB functions are intimately related, considering they are the different representations of a same function u. Actually, since

$$f(0,0) = \left[\frac{A_I(q_I)}{\det\sqrt{I - i\nabla^2 S(q_I)}} + \mathcal{O}(\epsilon)\right] \exp\left\{\frac{i}{\epsilon}\left(S_I(q_I) - \frac{p_I q_I}{2}\right)\right\},$$

applying Theorem 4.1 we have

$$f(0,t) = [A(t;z_I) + \mathcal{O}(\epsilon)] \exp\left(\frac{iS(t;z_I)}{\epsilon}\right),$$

where A and S solve the following ODEs:

$$\dot{S} + \frac{[z,\dot{z}]}{2} = 0, \;\; S(0;z_I) = S_I(q_I) - \frac{p_I q_I}{2}, \tag{5.11}$$

$$\dot{A} + \frac{\mathrm{tr}(Q^* \nabla^2 H(z)(iI - J)Q)}{2} A = 0, \;\; A(0;z_I) = \frac{A_I(q_I)}{\det\sqrt{I - i\nabla^2 S(q_I)}}. \tag{5.12}$$

The term $H(z)$ has been removed from (5.11) since $H(z) \equiv 0$ when $z \in \Lambda$. By Theorem 3.2 and formula (3.10), a coherent state integral approximation is then obtained as

$$\boxed{u(x) = (2\pi\epsilon)^{-\frac{N}{2}} \int_\Lambda A \exp\left(\frac{iS}{\epsilon}\right) [\rho_\epsilon(-z)\phi](x) d\mathrm{vol},} \tag{5.13}$$

where ϕ is defined by (2.4).

One might immediately point out shortcomings and even mistakes in the above derivation. The simpleness assumption on Λ is too restrictive and actually meaningless. More seriously, it is impossible for the global solution to have a compact support since the wave information will definitely propagate to infinity. However, the formula (5.13) itself is obviously well-defined, no matter how the Lagrangian manifold Λ is located in the phase space. Is it possible that based on a wrong reasoning, we have arrived at a correct solution? This is indeed the case, as shown in the next subsection.

5.3. Justification of the global asymptotic solution. Let us introduce an operator $\mathcal{K}$, analogous to the Maslov's canonical operator [28], as

$$[\mathcal{K}A](x)=(2\pi\epsilon)^{-\frac{N}{2}}\int_{\Lambda}A\exp\left(\frac{iS}{\epsilon}\right)[\rho_\epsilon(-z)\phi](x)d\mathrm{vol},\ \forall A\in\mathcal{C}^\infty(\Lambda),$$

where S is determined by (5.11) and ϕ is defined by (2.4).

THEOREM 5.1. *For any $A\in\mathcal{C}^\infty(\Lambda)$, we have*

$$H(W)\mathcal{K}A=(-i\epsilon)\mathcal{K}L(A)+\mathcal{O}(\epsilon^2),$$

where

$$L(A)=\dot{A}+\frac{\mathrm{tr}(Q^*\nabla^2H(z)(iI-J)Q)}{2}A.$$

The derivative is taken with respect to the variable t.

Proof. We first prove this theorem in the case that $\mathrm{supp}A$ is simple. Suppose $z=z(t)$ is a trajectory on $\mathrm{supp}A$. Set

$$f=\nu_\epsilon(z,U)\mathcal{K}A.$$

By (5.8), it holds that

$$\begin{aligned}\nu_\epsilon(z,U)H(W)\mathcal{K}A&=H(R_UW+z)\nu_\epsilon(z,U)\mathcal{K}A=H(R_UW+z)f\\&=\Big[H(R_UW+z)-R_UW\cdot\nabla H(z)+\frac{[z,\dot{z}]}{2}\\&\qquad-\frac{1}{2}(XT_IX+\epsilon^2DT_ID-\epsilon\mathrm{tr}T_I)-\epsilon XT_RD+(-i\epsilon)\partial_t\Big]f.\end{aligned}$$

This is a WKB function of VDP, since f is by Theorem 3.3. Considering that S solves (5.11), by Remark 4.1 and Lemma 3.3 we have

$$\begin{aligned}(\nu_\epsilon(z,U)H(W)\mathcal{K}A)_{x=0}&=\big(\square_0|_{x=0}+(-i\epsilon)\square_1|_{x=0}+\mathcal{O}(\epsilon^2)\big)\exp\left(\frac{iS}{\epsilon}\right)\\&=\big((-i\epsilon)L(A)+\mathcal{O}(\epsilon^2)\big)\exp\left(\frac{iS}{\epsilon}\right).\end{aligned}$$

Thus,

$$\begin{aligned}H(W)\mathcal{K}A&=(2\pi\epsilon)^{-\frac{N}{2}}\int_{\Lambda}(-i\epsilon)L(A)\exp\left(\frac{iS}{\epsilon}\right)[\rho_\epsilon(-z)\phi](x)d\mathrm{vol}+\mathcal{O}(\epsilon^2)\\&=(-i\epsilon)\mathcal{K}L(A)+\mathcal{O}(\epsilon^2).\end{aligned}$$

In the general case, suppose $\{e_j\}$ is a partition of unity of Λ such that $\mathrm{supp}e_j$ is simple. Then

$$H(W)\mathcal{K}A=\sum_j H(W)\mathcal{K}(e_jA)=(-i\epsilon)\sum_j\mathcal{K}L(e_jA)+\mathcal{O}(\epsilon^2)=(-i\epsilon)\mathcal{K}L(A)+\mathcal{O}(\epsilon^2).$$

The proof thus finishes. □

The above theorem implies that if we determine A by (5.12), then

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

which means that $\mathcal{K}A$ is a global asymptotic solution with accuracy to $\mathcal{O}(\epsilon)$. The formula (5.13) gives a first order global asymptotic solution of the wave equation (5.1), where A and S are determined by the following ODE system:

$$\dot{z} = J\nabla H(z),\ z(0) = (q_I, p_I), \tag{5.14}$$

$$\dot{C} = J\nabla^2 H(z)C,\ C(0) = \begin{pmatrix} I \\ \nabla^2 S(q_I) \end{pmatrix},\ C = QP, \tag{5.15}$$

$$\dot{S} + \frac{[z,\dot{z}]}{2} = 0,\ S(0) = S_I(q_I) - \frac{p_I q_I}{2}, \tag{5.16}$$

$$\dot{A} + \frac{\operatorname{tr}(Q^* \nabla^2 H(z)(iI - J)Q)}{2} A = 0,\ A(0) = \frac{A_I(q_I)}{\det\sqrt{I - i\nabla^2 S(q_I)}}. \tag{5.17}$$

The initial data $\nabla^2 S(q_I)$ is determined by (5.7).

6. Numerical tests

In this section, we report some numerical tests on Schrödinger and Helmholtz equations. The ODE systems (4.21)-(4.24) and (5.14)-(5.17) are integrated out with the "DIVPRK" subroutine in the IMSL library, which uses the Runge-Kutta-Verner fifth-order and sixth-order methods. The tolerance for error control is set as 10^{-14}, thus the computed quantities can be considered exact from the numerical point of view, if the evolution time is not too large. It should be pointed out that for long time evolutions, the symplectic integrator or spectral integrator would be more appropriate. However, we will not pursue this point in this paper.

We need also to approximate the continuous integrals in (4.20) and (5.13). For the Schrödinger equation, the initial position coordinate q_I is also a global coordinate of the Lagrangian manifold $\Lambda = \Lambda(t)$. In this case, we have

$$d\text{vol} = \det P dq_I.$$

The coordinate q_I is then discretized with the trapezoidal rule. Things become a little more complicated for the general scalar wave equation, since a global coordinate generally may not exist for the initial Lagrangian manifold Λ^I_{N-1}; see (5.5). However, if such a coordinate y exists, then (y,t) forms a global coordinate of the Lagrangian manifold Λ in (5.13). In this case, we have

$$d\text{vol} = \det P \left| \det \frac{\partial q}{\partial(y,t)} \bigg|_{q=q_I} \right| dydt.$$

The coordinate (y,t) is then discretized with the trapezoidal rule. Our numerical tests showed that in both cases, it suffices for the integration step size to be of order $\mathcal{O}(\sqrt{\epsilon})$ in each direction to guarantee a first order approximation to the continuous integrals. A numerical analysis will be reported in a forthcoming paper. In the general case when no global coordinate is available, we first resort to the partition of unity to decompose the initial Lagrangian manifold into a finite union of small patches, such that each patch admits a global coordinate. After deriving the wave solution for each patch, we sum up the wave fields together to form a global first order approximation.

In the following numerical tests, we indicate u_{CS} as the semi-classical approximate solution by the global geometrical optics method, and u either the exact or reference solution of the governing wave equation.

6.1. One-D Schrödinger equation. In this part, we consider the one-dimensional Schrödinger equation

$$i\epsilon\partial_t u = -\frac{\epsilon^2}{2}u_{xx} + V(x)u,\ u(x,0) = u_I(x).$$

As a first numerical example, we set

$$u_I(x) = \exp\left(-\frac{x^2}{2}\right)\exp\left(\frac{i\exp\left(-x^2\right)}{\epsilon}\right),\ V(x) = 0.$$

For any time $t \geq 0$, the displaced Lagrangian manifold associated with the initial WKB function is

$$\Lambda(t) = \{(x - 2txe^{-x^2}, -2xe^{-x^2}) \mid x \in \mathbb{R}\},$$

which is actually a curve in the two-dimensional phase space. Caustics will form at $t = 0.5$. At $t = 1.0$, there are two caustic points $x = \pm 0.285121\cdots$. Figure 6.1 shows the Lagrangian manifold at different times.

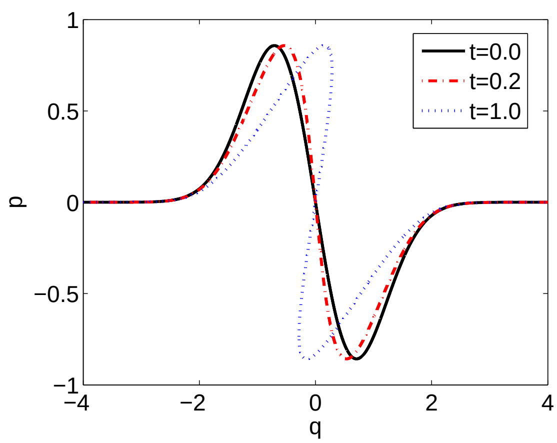

FIG. 6.1. *Lagrangian manifolds at different times in the phase space.*

The reference wave solution u is computed by the Fourier transform method. The computational domain is truncated into $[-10, 10]$, and 2^{18} grid points are used. Figure 6.2 compares the Gaussian coherent state approximation with the reference solution at $t = 1.0$ when $\epsilon = 2^{-8}$ is only moderately small. Even around caustics, they match very well.

Figure 6.3 depicts the relative errors at $t = 0.2$ and $t = 1.0$. A first order asymptotic convergence rate can be observed not only for the L^2-error, but for the L^∞-error also. Figure 6.4 shows the magnitude of error functions. The maximum error actually appears around caustics. It is not halved when ϵ goes from 2^{-15} to 2^{-16}. However, since the amplitude of wave solution increases at caustics, the relative error is indeed halved, as shown in figure 6.3.

Next we consider the Schrödinger equation with a variable potential. The initial data and the potential are set as

$$u_I(x) = \exp\left(-25x^2\right)\exp\left(\frac{i\cos(\pi x)}{\pi\epsilon}\right),\ V(x) = 2x^3.$$

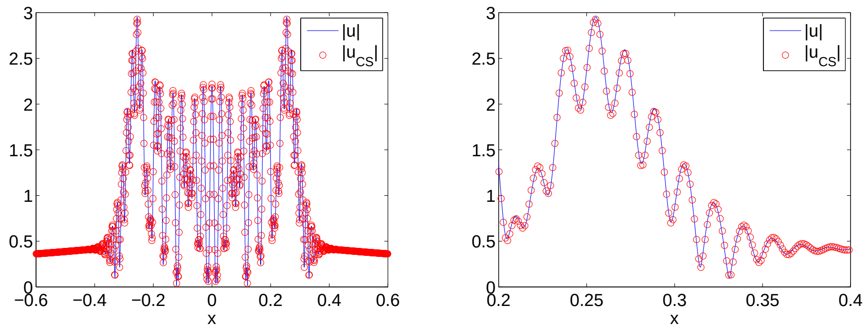

FIG. 6.2. *Comparison of wave amplitude.* $\epsilon = 2^{-8}$ *and* $t = 1.0$. *The right is zoomed in.*

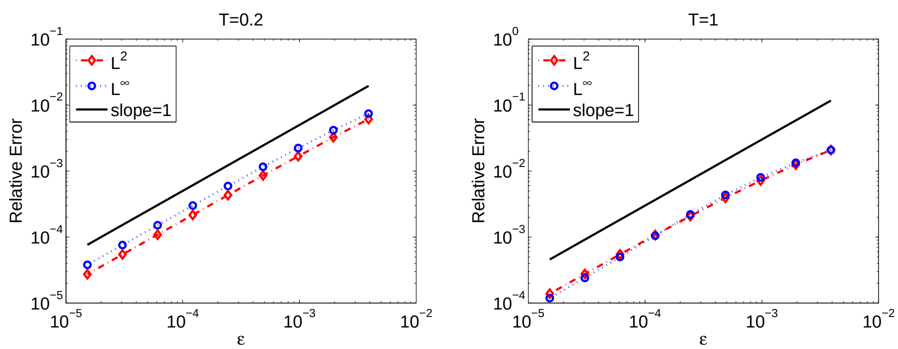

FIG. 6.3. *Relative error of coherent state approximation* $\frac{\|u-u_{CS}\|_\beta}{\|u\|_\beta}$ *with* $\beta = 2, \infty$. *Left:* $t = 0.2$. *Right:* $t = 1.0$.

This numerical setting has been applied in [19] to analyze the accuracy of the Lagrangian Gaussian beam approach. We compute the reference solution with the time-splitting spectral method [2] by truncating the computational domain into $[-2,2]$ and using 2^{18} grid points. The time step is set as 10^{-4}. There are two caustic points with significant energy when $t = 0.5$. In the left of figure 6.5 we compare the coherent state approximation with the reference solution, and in the right we depict the relative errors. A first order convergence rate is observed in both the L^2 norm and the L^∞ norm. For this example, the maximum error is attained around caustics only after ϵ becomes sufficiently small; see figure 6.6.

6.2. Two-D Schrödinger equation. We consider the two-dimensional free Schrödinger equation

$$i\epsilon\partial_t u = -\frac{\epsilon^2}{2}\Delta u,\ u(x,0) = A_I(x)\exp\left(\frac{iS_I(x)}{\epsilon}\right),$$

with

$$A_I(x) = \exp(-25x^2),\ S_I(x) = -\frac{1}{5}\log(4\cosh(5x_1)\cosh(5x_2)).$$

This numerical example has been used in [20] to demonstrate the performance of the level set method with phase shift correction. At $T = 0.5$, zigzag-type caustics

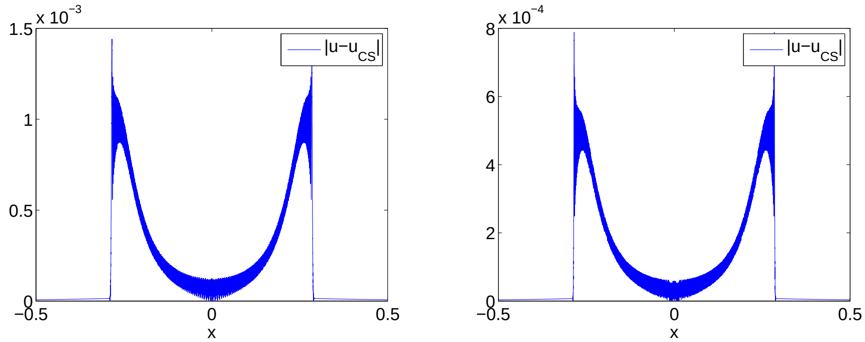

FIG. 6.4. *Error plot at* $t=1.0$. *Left:* $\epsilon=2^{-15}$. *Right:* $\epsilon=2^{-16}$.

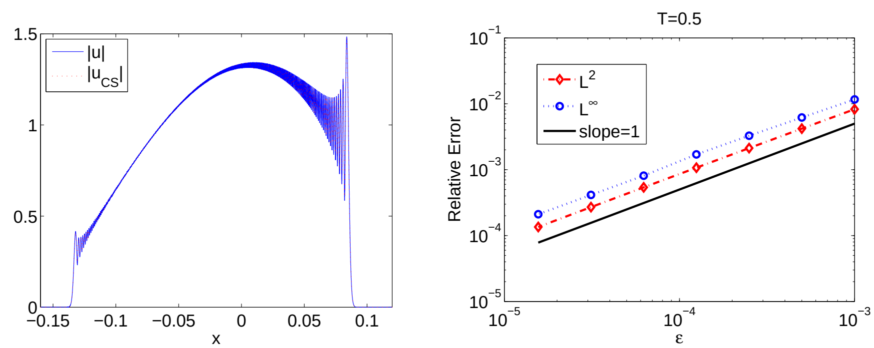

FIG. 6.5. $\epsilon=\frac{1}{16000}$, $t=0.5$. *Left: Comparison between the exact solution and the coherent state approximation. Right: Relative error of coherent state approximation* $\frac{\|u-u_{CS}\|_\beta}{\|u\|_\beta}$ *with* $\beta=2,\infty$.

will develop in both the x_1- and x_2-directions; see the illustrations in [20]. The left plot of figure 6.7 shows the wave amplitude of the reference solution derived by the Fourier transform method, while the right illustrates the approximate wave field by the coherent state approximation. Their difference is shown in the left of figure 6.8, and the relative maximum errors are shown in the right. A first order asymptotic convergence rate is clearly observed.

6.3. One-D Helmholtz equation. In this part, we consider the one-dimensional Helmholtz equation

$$-\epsilon^2 u_{xx} + \mathrm{erf}(x)u = 0.$$

Here erf is the error function, as shown in the left of figure 6.9. The Hamiltonian function is $H(x,\xi)=\xi^2+\mathrm{erf}(x)$. Note that for this example, the governing equation is only of hyperbolic-type on the negative real axis. Suppose the incident wave is right-going, i.e.,

$$u_I(x)=\exp\left(\frac{ix}{\epsilon}\right).$$

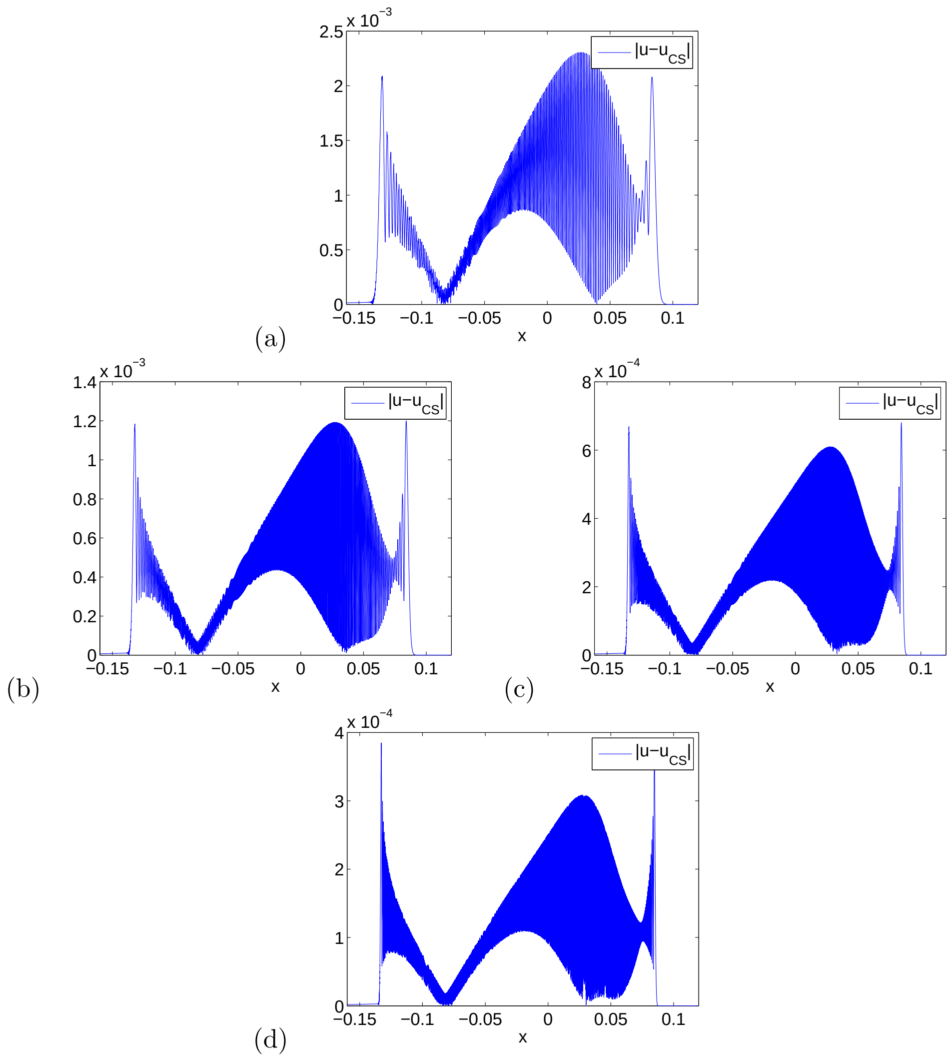

Fig. 6.6. *Error plot at* $t = 0.5$. *(a):* $\epsilon = \frac{1}{8000}$. *(b):* $\epsilon = \frac{1}{16000}$. *(c):* $\epsilon = \frac{1}{32000}$. *(d):* $\epsilon = \frac{1}{64000}$.

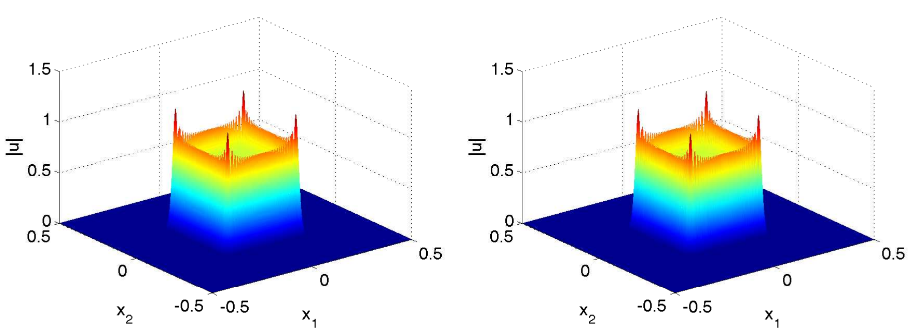

Fig. 6.7. $\epsilon = 0.001$. *Left: Reference solution. Right: Numerical solution.*

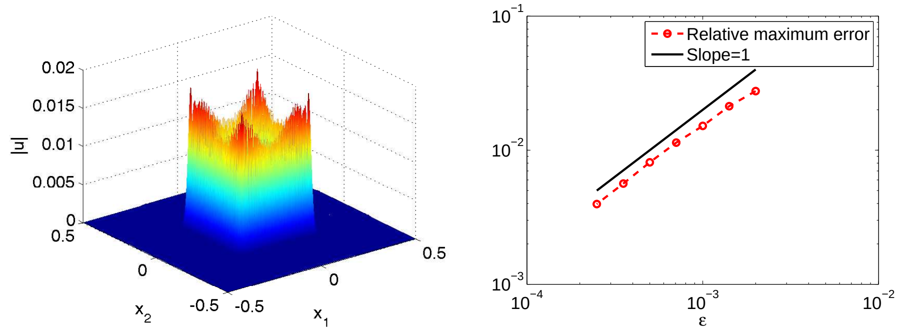

Fig. 6.8. *Left: Magnitude of error function with* $\epsilon = 0.001$. *Right: Convergence rate.*

The Lagrangian curve associated with this right-going wave is illustrated in the right of figure 6.9. The reference solution is derived by solving the truncated domain problem

$$
\begin{aligned}
&-\epsilon^2 u_{xx} + \mathrm{erf}(x) u = 0, \ x \in (x_l, x_r), \\
&u_x + \frac{iu}{\epsilon} = \frac{2i}{\epsilon} u_I, \ x = x_l, \\
&u_x + \frac{u}{\epsilon} = 0, \ x = x_r.
\end{aligned}
$$

In the numerical experiment, we set $[x_l, x_r] = [-5, 1]$ and employ the finite element method with sufficiently refined meshes. As illustrated in the right of figure 6.9, there exists one caustic point $x = 0$. The typical solution behavior is shown in the left of figure 6.10, and the relative maximum errors are plotted in the right. Again a first order asymptotic rate is observed.

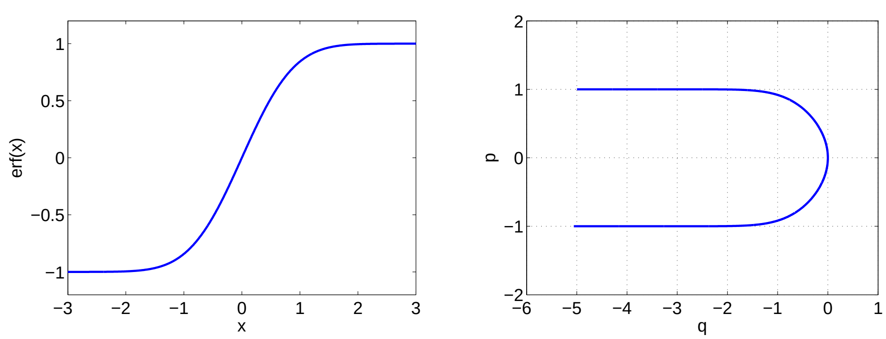

Fig. 6.9. *One-D Helmholtz equation. Left: error function. Right: Lagrangian curve.*

6.4. Two-D Helmholtz equation. In this subsection we consider the two-dimensional Helmholtz equation

$$\Delta u + k^2 u = 0.$$

The Hamiltonian function is

$$H(x, \xi) = \xi^2 - 1.$$

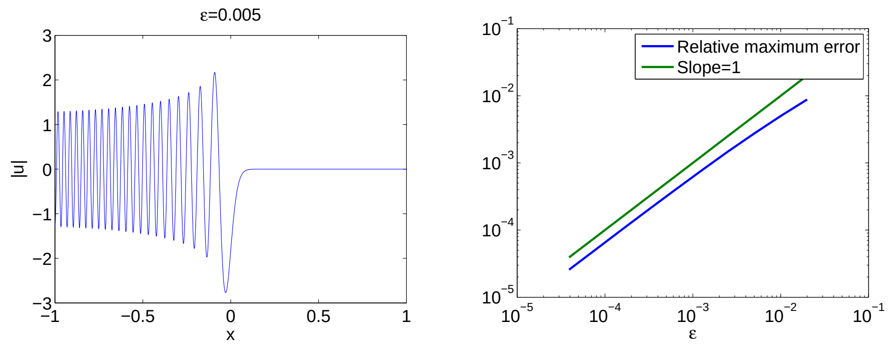

FIG. 6.10. *Left: reference solution with* $\epsilon = 0.005$. *Right: relative maximum errors.*

The Cauchy data is specified on the unit circle as $u_I = 1$, and we assume the wave propagates into the interior of this unit circle. It is straightforward to check that the incident wave field is actually

$$u_{inc} = \frac{H_0^{(2)}(kr)}{H_0^{(2)}(k)},$$

and the exact solution of the Helmholtz equation is simply

$$u = \frac{2J_0(kr)}{H_0^{(2)}(k)}.$$

The left of figure 6.11 shows the wave amplitude when $k = 100$. A focal point forms at the origin. The right of figure 6.11 plots the relative maximum errors. A first order asymptotic rate is obviously observed.

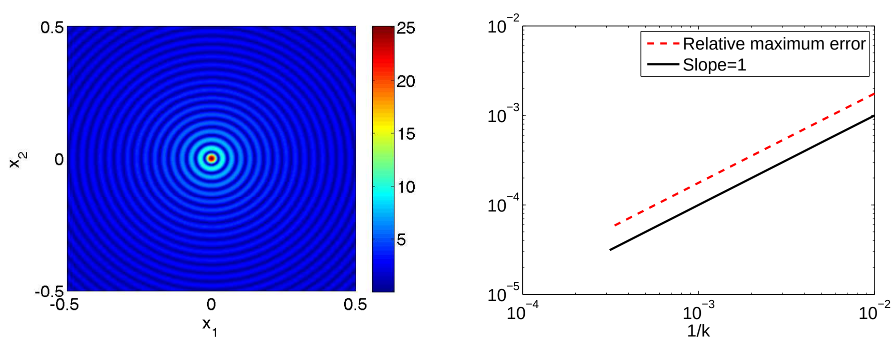

FIG. 6.11. *Left: solution amplitude with* $k = 100$. *Right: error plot.*

As a last numerical test, we consider the Helmholtz equation

$$\Delta u + k^2 n u = 0$$

with a variable index of refraction ($\sqrt{n}$) as

$$n(x) = 1 + \frac{1}{2}\exp\left(-50x_1^2 - 25x_2^2\right).$$

The plane wave $u_I(x)=\exp(ikx_1)$ travels from left to right. We illustrate the ray curves in the left of figure 6.12, from which a cusp caustic can be observed. The right of figure 6.12 illustrates the wave amplitude of the solution with $k=1000$, which is obtained by the proposed method.

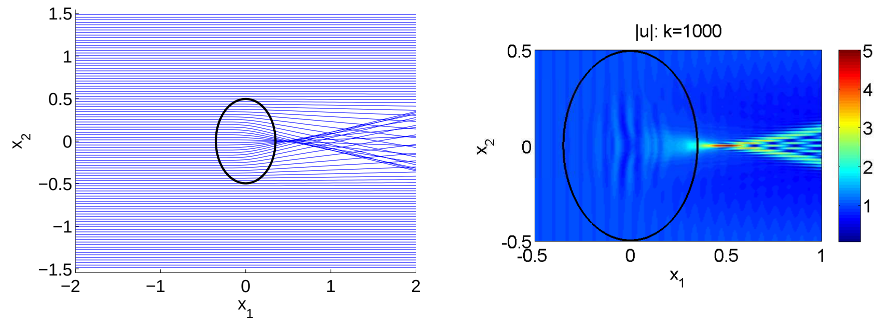

Fig. 6.12. *Plane wave diffracted by an elliptic lens. Left: ray curves. Right: wave amplitude.*

7. Conclusion

In this paper, we developed a new approach called the global geometrical optics method for the scalar linear wave equations in the high-frequency regime. We first considered the Schrödinger equations. The basic idea of the proposed method is to express the governing wave equation under a moving frame determined by the underlying Hamiltonian flow. The benefit of this treatment is such that a global in time WKB-ansatz solution is valid for an arbitrary but fixed evolution time, if the initial WKB data is simple enough. In addition, the phase and the amplitude at the origin $x=0$ can be determined merely through two simple ODEs. By varying the starting point of the moving frame, we form a WKB-type function only defined on the Lagrangian manifold, from which the wave field can be retrieved with accuracy to first order. For more general initial WKB data, we employed the idea of partition of unity. However, the final semi-classical approximation does not rely on the specific choice of partition of unity. We also extended the global geometrical optics method to the general scalar wave equation with Cauchy data. We define an analogy of Maslov's canonical operator, from which a globally valid semi-classical approximation with Gaussian coherent states was proposed.

We would like to make a comparison between the proposed method and the Gaussian beam approach. A common point shared by these two methods is that the approximation is formulated as some sort of stationary path integral in the phase space, and the numerical implementation can be realized by the method of characteristics. Besides, both methods can solve the caustic problem to some extent suffered by the classical geometrical optics method. However, the accuracy of thawed-type Gaussian beam approximation deteriorates around caustics, as exposed in [19]. The frozen-type Gaussian beam approach such as the Herman-Kluk approximation [13, 21] is formulated as an integral in much higher dimensions. In d dimensions, the Herman-Kluk approximation is defined on a $3d$ manifold, though some improved version [32] reduces the integrating space to $2d$ dimensions. Comparatively, our approach (see (4.20) or (5.13)) presents a uniform first order approximation, and is only defined on a d-dimensional manifold. We should be frank, however, that the proposed method is restricted to the initial data in the WKB form (or a finite sum of them), while other approximations are applicable to much more general initial data. In addition, it is not clear for the moment whether or not the semi-classical approximation (4.20) can be straightforwardly derived from either the Herman-Kluk approximation or some of its analogues by a proper stationary phase analysis.

Acknowledgments. This work is supported by the National Natural Science Foundation grant No. 10971115, and Tsinghua University Initiative Scientific Research Program.

REFERENCES

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

[2] W. Bao, S. Jin, and P.A. Markowich, *On time-splitting spectral approximations for the Schrödinger equation in the semiclassical regime*, J. Comput. Phys., 175, 487–524, 2002.

[3] Y. Brenier and L. Corrias, *A kinetic formulation for multibranch entropy solutions of scalar conservation laws*, Ann. IHP Nonlin. Anal., 15, 169–190, 1998.

[4] S. Bougacha, J.L. Akian, and R. Alexandre, *Gaussian beams summation for the wave equation in a convex domain*, Commun. Math. Sci., 7(4), 973–1008, 2009.

[5] A. Bultheel and H. Martínez, *A Shattered Survey of the Fractional Fourier Transform*, TW Reports, K.U. Leuveven, TW337, 2002.

[6] L.T. Cheng, H.L. Liu, and S. Osher, *Computational high frequency wave propagation using the level set method with applications to the semiclassical limit of the Schrödinger equation*, Commun. Math. Sci., 1, 593–612, 2003.

[7] B. Enquist and O. Runborg, *Multi-phase computations in geometrical optics*, J. Comput. Appl. Math., 74, 175–192, 1996.

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

[9] P. Gerard, P.A. Markowich, N.J. Mauser, and F. Poupaud, *Homogenization limits and Wigner transforms*, Commun. Pure Appl. Math., 50, 323–379, 1997.

[10] L. Gosse, *Using K-branch entropy solutions for multivalued geometric optics computations*, J. Comput. Phys., 180, 155–182, 2002.

[11] L. Gosse, S. Jin, and X. Li, *On two moment systems for computing multiphase semiclassical limits of the Schrödinger equation*, Math. Model Meth. Appl. Sci., 13, 1689–1723, 2003.

[12] E.J. Heller, *Time-dependent approach to semiclassical dynamics*, J. Chem. Phys., 62, 1544, 1975.

[13] 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.

[14] L. Hörmander, *On the existence and the regularity of solutions of linear pseudo-differential equations*, L'Enseignement Mathématique, XVII, 99–163, 1971.

[15] S. Jin, *Recent computational methods for high frequency waves in heterogeneous media*, Industrial and Applied Mathematics in China, Ser. Contemp. Appl. Math. CAM, Higher Ed. Press, Beijing, 10, 49–64, 2009.

[16] S. Jin and X. Li, *Multiphase computation of the semiclassical limit of the Schrödinger equation and related problems: Whitham vs. Wigner*, Physica D, 182, 46–85, 2003.

[17] S. Jin, H. Liu, S. Osher, and R. Tsai, *Computing multivalued physical observables for the semiclassical limit of the Schrödinger equation*, J. Comput. Phys., 205, 222–241, 2005.

[18] S. Jin and S. Osher, *A level set method for the computation of multivalued solutions to quasi-linear hyperbolic PDEs and Hamilton-Jacobi equations*, Commun. Math. Sci., 1(3), 575–591, 2003.

[19] 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.

[20] S. Jin and X. Yang, *Computation of the semiclassical limit of the Schrödinger equation with phase shift by a level set method*, J. Sci. Comput., 35, 144–169, 2008.

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

[22] J.B. Keller, *Corrected Bohr-Sommerfeld qunatum conditions for non-separable systems*, Ann. Phys., 4, 180–188, 1958.

[23] H. Liu, L.T. Cheng, and S. Osher, *A level set framework for capturing multi-valued solutions of nonlinear first-order equations*, J. Sci. Comput., 29, 353–373, 2006.

[24] H. Liu, S. Osher, and R. Tsai, *Multi-valued solution and level set methods in computational high frequency wave propagation*, Commun. Comput. Phys., 1, 765–804, 2006.

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

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

[27] H. Liu, O. Runborg, and N.M. Tanushev, *Error estimates for Gaussian beam superpositions*, Math. Comp., to appear.

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

[29] M. Motamed and O. Runborg, *Taylor expansion and discretization errors in Gaussian beam superposition*, Wave Motion, 2010.

[30] C. Paige and C. van Loan, *A Schur decomposition for Hamiltonian matrices*, Lin. Alg. Appl., 41, 11–32, 1981.

[31] J. Ralston, *Gaussian beams and the propagation of singularities*, Stud. Part. Diff. Eqs., MAA Stud. Math., Math. Assoc. America, Washington DC, 23, 206–248, 1982.

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

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

[34] C. Zheng, *Gaussian beam approach for the boundary value problem of high frequency Helmholtz equation*, Commun. Math. Sci., 8(4), 1041–1066, 2010.