Euro. Jnl of Applied Mathematics (2005), *vol.* 16, *pp.* 1–30.
doi:10.1017/S0956792505006303 Printed in the United Kingdom

On the evolution of the semi-classical Wigner function in higher dimensions

S. FILIPPAS [1,2] and G. N. MAKRAKIS [1,2]

[1] *Department of Applied Mathematics, University of Crete, 71409 Heraklion, Greece*
[2] *Institute of Applied and Computational Mathematics, FORTH, 71110 Heraklion, Greece*
Email: {filippas,makrakg}@tem.uoc.gr

(*Received 17 May 2004; revised 29 April 2005*)

The limit Wigner measure of a WKB function satisfies a simple transport equation in phase-space and is well suited for capturing oscillations at scale of order $O(\epsilon)$, but it fails, for instance, to provide the correct amplitude on caustics where different scales appear. We define the semi-classical Wigner function of an N-dimensional WKB function, as a suitable formal approximation of its scaled Wigner function. The semi-classical Wigner function is an oscillatory integral that provides an ϵ-dependent regularization of the limit Wigner measure, it obeys a transport-dispersive evolution law in phase space, and it is well defined even at simple caustics.

1 Introduction and preliminaries

If $\psi^\epsilon(\mathbf{x})$ is a family of functions that decay rapidly at infinity, and $\epsilon > 0$ is a small parameter, the scaled Wigner transform of ψ^ϵ is defined.

$$W^\epsilon(\mathbf{x},\mathbf{k}) = \frac{1}{(2\pi)^N}\int_{\mathbf{R}^N} e^{-i\mathbf{k}\cdot\mathbf{y}}\psi^\epsilon\left(\mathbf{x}+\frac{\epsilon\mathbf{y}}{2}\right)\overline{\psi^\epsilon}\left(\mathbf{x}-\frac{\epsilon\mathbf{y}}{2}\right)d\mathbf{y}, \qquad (\mathbf{x},\mathbf{k})\in\mathbf{R}^N\times\mathbf{R}^N, \tag{1.1}$$

where $\overline{\psi^\epsilon}(\mathbf{x})$ is the complex conjugate of ψ^ϵ. This is a real function in phase space and its k-integral gives the amplitude of $\psi^\epsilon(\mathbf{x})$,

$$\int_{\mathbf{R}^N} W^\epsilon(\mathbf{x},\mathbf{k})\,d\mathbf{k} = |\psi^\epsilon(\mathbf{x})|^2. \tag{1.2}$$

Hence, we may think of W^ϵ as wave number resolved energy density. This is not quite precise however, because W^ϵ is not in general positive, except when ψ^ϵ is a Gaussian function (see Lions & Paul [16]), but it always becomes positive in the high frequency limit. In particular, as the small parameter ϵ tends to zero, the Wigner function W^ϵ tends weakly to a positive measure W^0 called the limit Wigner measure [16]. One remarkable property of the Wigner transform is the following: If ψ^ϵ is taken to be the solution of a large class of homogenization problems for evolution equations, then the corresponding limit Wigner measure solves a simple transport equation in phase space. Once the limit Wigner measure is known one can recover useful information about the underlying function ψ^ϵ as $\epsilon\to 0$. This makes the Wigner transform a particularly powerful tool in the study of high frequency wave propagation problems (e.g. see the book by Tatarskii [23], and the expository paper by Papanicolaou & Ryzhik [19]).

To illustrate this fact we consider a concrete problem. It is well known that the paraxial approximation for several classical wave equations leads to the Cauchy problem for the time-dependent $(N+1)$-dimensional Schrödinger equation with fast space-time scales [14, 22]

$$i\epsilon\psi_t^\epsilon(\mathbf{x},t) = -\frac{\epsilon^2}{2}\Delta\psi^\epsilon(\mathbf{x},t) + V(\mathbf{x})\psi^\epsilon(\mathbf{x},t), \qquad \mathbf{x}\in\mathbf{R}^N, \tag{1.3}$$

and highly oscillatory initial data

$$\psi^\epsilon(\mathbf{x},0) = \psi_0^\epsilon(\mathbf{x}) = A_0(\mathbf{x})\exp\big(iS_0(\mathbf{x})/\epsilon\big). \tag{1.4}$$

Note that the small parameter ϵ appears in both the equation and the initial data. We are interested in the high frequency limit of (1.3), (1.4), that is, in the limit of ψ^ϵ as ϵ tends to zero.

In the standard high frequency approximation (WKB method) we look for an asymptotic solution of (1.3), (1.4), in the same form as the initial data

$$\psi^\epsilon(\mathbf{x},t) = A(\mathbf{x},t)\exp\big(iS(\mathbf{x},t)/\epsilon\big). \tag{1.5}$$

If we plug in this into (1.3) and equate powers of ϵ, we obtain evolution equations for the phase $S(\mathbf{x},t)$ and the amplitude $A(\mathbf{x},t)$,

$$S_t + \frac{1}{2}|\nabla_\mathbf{x}S|^2 + V(\mathbf{x}) = 0, \qquad S(\mathbf{x},0) = S_0(\mathbf{x}), \tag{1.6}$$

$$(A^2)_t + \nabla_\mathbf{x}\cdot(A^2\nabla_\mathbf{x}S) = 0, \qquad A(\mathbf{x},0) = A_0(\mathbf{x}). \tag{1.7}$$

Equation (1.6) is called the eikonal equation and (1.7) the transport equation. The solution of the eikonal equation can be constructed by the method of characteristics (rays). The characteristic curves $\overline{\mathbf{x}}(\mathbf{q},t)$ and $\overline{\mathbf{k}}(\mathbf{q},t) = S_x(\overline{\mathbf{x}}(\mathbf{q},t),t)$ are first obtained by solving the ODE system

$$\frac{d}{dt}\overline{\mathbf{x}} = \overline{\mathbf{k}}, \qquad \frac{d}{dt}\overline{\mathbf{k}} = -\nabla V(\overline{\mathbf{x}}), \tag{1.8}$$

with $\overline{\mathbf{x}}(0) = \mathbf{q}$ and $\overline{\mathbf{k}}(0) = \nabla S_0(\mathbf{q})$. Then the phase $S = S(\overline{\mathbf{x}}(\mathbf{q},t),t)$ is obtained by integrating the equation

$$\frac{dS}{dt} = S_t + |\nabla_\mathbf{x}S|^2 = \frac{|\overline{\mathbf{k}}|^2}{2} - V, \qquad S(\overline{\mathbf{x}},0) = S_0(\mathbf{q}),$$

along the characteristics $\overline{\mathbf{x}}(\mathbf{q},t)$.

Since (1.6) is a nonlinear equation, it has, in general, a smooth solution only up to some finite time t_c, when the rays cross each other and singularities are developed that are called *caustics*. At a caustic point the Jacobian $J(\mathbf{q},t)$ of the ray map $\mathbf{q}\mapsto\overline{\mathbf{x}}(\mathbf{q},t)$ is zero, and the amplitude becomes infinite. This follows from the formula

$$A(\mathbf{x},t) = A_0(\mathbf{q})J^{-1/2}(\mathbf{q},t), \tag{1.9}$$

which is obtained by integrating the transport equation over a ray tube and then using the divergence theorem [2].

Let us now see how one can treat this problem using the Wigner transform formalism. If we denote by $f^\epsilon(\mathbf{x},\mathbf{k},t)$ the Wigner transform of the solution $\psi^\epsilon(\mathbf{x},t)$ of (1.3), (1.4), then f^ϵ satisfies the *Wigner equation* ([16], [19], [9])

$$f^\epsilon_t(\mathbf{x},\mathbf{k},t) + \mathbf{k}\cdot\nabla_{\mathbf{x}} f^\epsilon(\mathbf{x},\mathbf{k},t) + \mathscr{L}_\epsilon f^\epsilon(\mathbf{x},\mathbf{k},t) = 0, \tag{1.10}$$

where the operator $\mathscr{L}_\epsilon$ is defined by the convolution with respect to the momentum $\mathbf{k}$,

$$\mathscr{L}_\epsilon f(\mathbf{x},\mathbf{k},t) = f(\mathbf{x},\mathbf{k},t) *_k \frac{i}{(2\pi\epsilon)^N}\int_{-\infty}^{\infty}\exp(-i\mathbf{k}\cdot\mathbf{y})\left(V\left(\mathbf{x}+\frac{\epsilon}{2}\mathbf{y}\right) - V\left(\mathbf{x}-\frac{\epsilon}{2}\mathbf{y}\right)\right)d\mathbf{y}. \tag{1.11}$$

Assuming that the potential V is smooth enough, we can expand $V(\mathbf{x}\pm\frac{\epsilon}{2}\mathbf{y})$ into Taylor series to get

$$\frac{1}{\epsilon}\left(V\left(\mathbf{x}+\frac{\epsilon}{2}\mathbf{y}\right) - V\left(\mathbf{x}-\frac{\epsilon}{2}\mathbf{y}\right)\right) = \mathbf{y}\cdot\nabla V(\mathbf{x}) + \sum_{m\geqslant 1}\left(\frac{\epsilon}{2}\right)^{2m}\sum_{|\alpha|=2m+1}\frac{\mathbf{y}^\alpha}{\alpha!}D^\alpha V(\mathbf{x}),$$

so that we can rewrite equation (1.10) as [24]

$$f^\epsilon_t + \mathbf{k}\cdot\nabla_{\mathbf{x}} f^\epsilon - \nabla V\cdot\nabla_{\mathbf{k}} f^\epsilon = \sum_{m\geqslant 1} c_m\epsilon^{2m}\sum_{|\alpha|=2m+1} D^\alpha V D^\alpha_{\mathbf{k}} f^\epsilon, \tag{1.12}$$

where $c_m = \frac{(-1)^m}{2^{2m}(2m+1)!}, m = 0,1,\ldots.$ The initial condition for (1.10) or (1.12) is the Wigner transform of the initial data (1.4) that we denote by f^ϵ_0,

$$f^\epsilon_0(\mathbf{x},\mathbf{k},t=0) = \frac{1}{(\epsilon\pi)^N}\int_{\mathbf{R}^N} A_0(\mathbf{x}+\boldsymbol{\sigma})A_0(\mathbf{x}-\boldsymbol{\sigma})e^{\frac{i}{\epsilon}(S_0(\mathbf{x}+\boldsymbol{\sigma})-S_0(\mathbf{x}-\boldsymbol{\sigma})-2\mathbf{k}\cdot\boldsymbol{\sigma})}\,d\boldsymbol{\sigma}. \tag{1.13}$$

In the formal limit $\epsilon = 0$, f^ϵ tends to the limit Wigner measure f^0 and the dispersive part of the Wigner equation, that is, the right hand side of (1.12), vanishes. It follows that f^0 satisfies the Liouville equation

$$f^0_t(\mathbf{x},\mathbf{k},t) + \mathbf{k}\cdot\nabla_{\mathbf{x}} f^0(\mathbf{x},\mathbf{k},t) - \nabla V(\mathbf{x})\cdot\nabla_{\mathbf{k}} f^0(\mathbf{x},\mathbf{k},t) = 0\,, \tag{1.14}$$

which is a simple transport equation in the phase space $\mathbf{R}^{2N}_{\mathbf{xk}}$. The initial condition for (1.14) is the (weak) limit of f^ϵ_0, as $\epsilon\to 0$, in (1.13),

$$f^0_0(\mathbf{x},\mathbf{k},t=0) = A_0^2(\mathbf{x})\delta(\mathbf{k}-\nabla S_0(\mathbf{x})). \tag{1.15}$$

It is not difficult to verify that the solution of (1.14), (1.15), in the time interval $[0,T]$, $T < t_c$, is given by

$$f^0(\mathbf{x},\mathbf{k},t) = A^2(\mathbf{x},t)\delta(\mathbf{k}-\nabla_{\mathbf{x}} S(\mathbf{x},t)), \tag{1.16}$$

where $S(\mathbf{x},t)$ and $A(\mathbf{x},t)$ are solutions of the eikonal (1.6) and transport (1.7) equations, respectively. Thus, from the limit Wigner measure we can recover the modulus of the amplitude A and the gradient of the phase S of the WKB approximation. In particular, we have

$$\int_{\mathbf{R}^N} f^0(\mathbf{x},\mathbf{k},t)\,d\mathbf{k} = A^2(\mathbf{x},t). \tag{1.17}$$

However, the Wigner approach has certain advantages over the WKB method since it provides flexibility to treat more general initial data; we refer elsewhere [21, 13, 8] for a thorough comparison of the two methods.

Besides its connection with classical wave propagation, equation (1.3)–(1.4) is the fundamental model of quantum mechanics [17], where ψ_ϵ is the probability density of the position of a particle with unit mass, and ϵ plays the role of the Planck constant. Whereas the limit Wigner measure is always positive and corresponds to the classical motion, the scaled Wigner function is highly oscillatory and sign changing, thus taking into account quantum interference and coherence phenomena. In this context, both the questions of approximation of the scaled Wigner function as well as its evolution in time, are of high importance. See the references [3, 4, 18, 24] and more recently [20, 11] and references therein.

Our motivation for the present work comes from the question of what can be said about the solution f^ϵ of the Wigner equation (1.12) with WKB initial data (1.13) when ϵ is small but not zero. Formally speaking, equation (1.12) is a transport-dispersive equation infinitely singular (as ϵ goes to zero) and therefore serious analytical as well as numerical difficulties are anticipated. On the other hand, even simple one dimensional examples [8], show that f^ϵ has a very complicated structure, which is expressed through generalized Airy functions. Therefore, we search for a suitable approximant of f^ϵ which, to some extend, captures some of the basic features of f^ϵ.

Although in this paper we are interested in higher dimensions ($N \geqslant 2$), we recall here some observations for the one-dimensional case ($N = 1$) considered in [8], to give some insight about the solution f^ϵ of the Wigner equation. Let us consider the equation (1.3)–(1.4) for $V \equiv 0$, with initial data $A_0(x) = 1$ and $S_0(x) = -x^3/3$. This special case can be worked out explicitly. The initial scaled Wigner function is given by the Airy function

$$f_0^\epsilon(x,k,t=0) = \frac{2^{\frac{2}{3}}}{\epsilon^{\frac{2}{3}}}\,\mathrm{Ai}\left(\frac{2^{\frac{2}{3}}}{\epsilon^{\frac{2}{3}}}(k+x^2)\right), \tag{1.18}$$

whereas the initial limit Wigner measure, that is the limit of f_0^ϵ as $\epsilon \to 0$, is $f_0^0(x,k,t=0) = \delta(k+x^2)$. If one integrates f_0^ϵ with respect to k and uses the identity $\int_{\mathbf{R}} \mathrm{Ai}(z)dz = 1$, then one recovers the amplitude ($A_0 = 1$). Alternatively, one can integrate f_0^0 to reach the same result. The observation here is that although f_0^ϵ is "rich" and has a complex structure (the Airy oscillations), it conveys *no extra* information – compared to f_0^0 – as far as the amplitude is concerned. Let us also observe that f_0^ϵ provides the *proper regularization* of f_0^0. Loosely speaking, before f_0^0 becomes a Dirac mass in the limit $\epsilon \to 0$, it was an Airy function and not, for instance, an ϵ-sequence of "top-hat"-functions. This is what always happens in the single phase case, but also in the multi-phase case away from caustics.

On a caustic however, this "equivalence" of f^ϵ and f^0 is lost, in the sense that f^0 is unable to provide the amplitude but f^ϵ can. Caustic points are the *only* points where "keeping the ϵ" is crucial for recovering the correct amplitude. In our specific example, a fold caustic appears at (x_f, t_f), $t_f > 0$; see Filippas & Makrakis [8, §4.1] for detailed calculations. Although f^0 is a well defined Delta function on $\mathbf{R}_x \times \mathbf{R}_k$ for any $t > 0$, its restriction at the point (x_f, t_f), which is formally given by $f^0(x_f, k, t_f) = \delta(t_f^2(k - k_f)^2)$, is not a well defined distribution in $\mathbf{R}_k$; see Lax [15, p. 547]. Therefore, the projection identity (1.17) cannot be used for $f^0(x_f, \cdot, t_f)$ for computing the amplitude. Instead, one has to use the scaled Wigner function

$$f^\epsilon(x_f, k, t_f) = \frac{2^{\frac{2}{3}}}{\epsilon^{\frac{2}{3}}} \mathrm{Ai}\left(\frac{2^{\frac{2}{3}}}{\epsilon^{\frac{2}{3}}} t_f^2 (k - k_f)^2\right),$$

and the projection formula (1.2) in order to recover the precise amplitude.

One way to see why an "object" like $\delta(y^2)$ is not well defined, is by noting that if $\phi_\epsilon(y)$ is a regularization of the Dirac mass $\delta(y)$ (that is, a sequence of smooth functions that tends to $\delta(y)$ as ϵ tends to 0), then the limit of $\phi_\epsilon(y^2)$ is not uniquely defined but depends on *the sequence itself*. Thus, at a caustic point, it is of high importance that the proper regularization of f^0 is the Airy and not, say, an ϵ-sequence of "top-hat"-functions. Consequently, when working with f^0, an important information – where f^0 comes from – is lost, and f^0, which is rather efficient in describing quantities that involve oscillations at a scale of order $O(\epsilon)$, fails to do so on caustics, where different scales (like $O(\epsilon^{\frac{2}{3}})$ in our example) develop.

The above discussion suggests that an essential characteristic that a reasonable approximant of f^ϵ should have, is to provide the proper regularization of the limit Wigner measure.

In this work we consider equation (1.12) with WKB initial data (1.13), and for the single-phase case, that is, for times prior to t_c, we obtain an approximant $\widetilde{f}^\epsilon$ of f^ϵ. This approximant is expressed in terms of a certain oscillatory integral P_N (see (5.3) and (2.8) below) and it has the following basic features:

(i) It is in agreement with the (scaled Wigner function of the) WKB solution, in a sense that we will make precise later;

(ii) it provides the proper regularization of the limit Wigner Dirac mass;

(iii) it obeys a transport–dispersive evolution law which is in agreement with the transport–dispersive character of the Wigner equation (1.12)–(1.13); and

(iv) it is well defined even on simple caustics at time t_c. At such points the projection formula (1.17) is inapplicable, but the integration of $\widetilde{f}^\epsilon$ with respect to $\mathbf{k}$ is meaningful. Although an approximation result relating the integral of $\widetilde{f}^\epsilon$ and the exact amplitude $|\psi^\epsilon(\mathbf{x})|^2$ is still missing, this observation could be useful for numerical calculations.

Our strategy towards the construction of $\widetilde{f}^\epsilon$, is the same as in the one dimensional case [8], and can be roughly described as follows. Departing from the WKB solution (1.5), we first define by (2.10) the semi-classical Wigner function $\widetilde{W}^\epsilon$, as a formal approximant of the scaled Wigner function of a WKB function at a fixed time t. To this end, we introduce and study a suitable N-dimensional phase integral that reduces to the Airy function when $N = 1$. $\widetilde{W}^\epsilon$ in fact provides the proper (P_N) regularization of the limit Wigner Dirac mass. We then define $\widetilde{f}^\epsilon$ as the evolution of the initial semi-classical Wigner function ($\widetilde{W}_0^\epsilon$) under an appropriate evolution law (see (5.2)). The key point here is the derivation of the evolution law. This is done by requiring that the evolved semi-classical Wigner function ($\widetilde{f^{\epsilon}}$) be in agreement with the approximant $\widetilde{W^{\epsilon}}$ for times $t < t_c$. It then follows that, $\widetilde{f^{\epsilon}}$ is given by the convolution of $\widetilde{W_0^{\epsilon}}$ with a suitable kernel $\tilde{G}^{\epsilon}$.

It is important to note that although the construction of $\widetilde{f^{\epsilon}}$ is based on the WKB solution, the evolution law for $\widetilde{f^{\epsilon}}$ incorporates the expected structure of f^{ϵ}, and provides some insight in what one might expect from the Wigner equation (1.12)–(1.13). In particular, $\widetilde{f^{\epsilon}}$ combines the transport of the initial data along the bicharacteristics, as it is required by the Liouville part of the Wigner equation, with an ϵ-dependent dispersion mechanism as required by the dispersive right-hand side of (1.12). The dispersion mechanism is realized as a convolution of the transported initial data with the kernel $\tilde{G}^{\epsilon}$. As $\epsilon \to 0$ then $\widetilde{f^{\epsilon}}$ tends to the limit Wigner Dirac mass, $\tilde{G}^{\epsilon}$ tends to a Dirac mass, and we recover the Liouville equation that the limit Wigner measure obeys.

The paper is organized as follows. In § 2 we define the semi-classical Wigner function $\widetilde{W^{\epsilon}}$, as a formal approximant of the scaled Wigner W^{ϵ} of a WKB function. To this end we need to introduce a suitable phase integral P_N and study some of its properties; this is done in appendix A. We note that the semi-classical Wigner that corresponds to the WKB solution (1.5), at each time $t < t_c$, depends on the vector $\mathbf{k} - \nabla_{\mathbf{x}} S(\mathbf{x}, t)$ and the third-derivative tensor $S_{x_i,x_j,x_k}(\mathbf{x}, t)$.

In § 3, we derive the evolution law of $(\mathbf{k} - \nabla_{\mathbf{x}} S(\mathbf{x}, t))$ and $S_{x_i,x_j,x_k}(\mathbf{x}, t)$. The most technical part, concerned with the study of a matrix differential equation, is moved to Appendix B. These results will be used in § 4 and 5.

§ 4 serves as an introduction to § 5. We solve the Liouville equation satisfied by the limit Wigner measure, and we motivate the definition of $\widetilde{f^{\epsilon}}$ through a suitable evolution law that will be given in the next section. § 5 contains the main result of this work, that is, the derivation of the appropriate evolution law for the semi-classical Wigner function, and the definition of $\widetilde{f^{\epsilon}}$.

In the final § 6, we study a simple two dimensional example, involving an elliptic umbilic caustic. In this example, the exact f^{ϵ} is computed explicitly at any time $t \geqslant 0$. We show how the projection formula (1.2) can be used on a caustic to give the precise amplitude, in conjunction with suitable phase-space identities derived in Berry & Wright [5]. A particular case of such an identity is recalled in Appendix C.

2 The semi-classical Wigner of a WKB function

We recall that the scaled Wigner transform of the function $\psi^{\epsilon}(\mathbf{x})$ is defined by

$$\begin{aligned} W^{\epsilon}(\mathbf{x}, \mathbf{k}) &= \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{-i\mathbf{k}\cdot\mathbf{y}} \psi^{\epsilon}\left(\mathbf{x} + \frac{\epsilon \mathbf{y}}{2}\right) \overline{\psi^{\epsilon}}\left(\mathbf{x} - \frac{\epsilon \mathbf{y}}{2}\right) d\mathbf{y} \\ &= \frac{1}{(\epsilon\pi)^N} \int_{\mathbf{R}^N} e^{-\frac{2i}{\epsilon}\mathbf{k}\cdot\boldsymbol{\sigma}} \psi^{\epsilon}(\mathbf{x} + \boldsymbol{\sigma}) \overline{\psi^{\epsilon}}(\mathbf{x} - \boldsymbol{\sigma}) \, d\boldsymbol{\sigma}. \end{aligned} \tag{2.1}$$

Therefore the scaled Wigner transform of a WKB function $\psi^{\epsilon}(\mathbf{x}) = A(\mathbf{x}) e^{i\frac{S(\mathbf{x})}{\epsilon}}$ is given by

$$W^{\epsilon}(\mathbf{x}, \mathbf{k}) = \frac{1}{(\epsilon\pi)^N} \int_{\mathbf{R}^N} D(\mathbf{x}, \boldsymbol{\sigma}) e^{i\frac{F(\mathbf{x}, \mathbf{k}; \boldsymbol{\sigma})}{\epsilon}} \, d\boldsymbol{\sigma}, \tag{2.2}$$

with

$$D(\mathbf{x};\boldsymbol{\sigma}) := A(\mathbf{x}+\boldsymbol{\sigma})A(\mathbf{x}-\boldsymbol{\sigma}), \tag{2.3}$$

$$F(\mathbf{x},\mathbf{k};\boldsymbol{\sigma}) := S(\mathbf{x}+\boldsymbol{\sigma}) - S(\mathbf{x}-\boldsymbol{\sigma}) - 2\mathbf{k}\cdot\boldsymbol{\sigma}. \tag{2.4}$$

We will define the semi-classical Wigner function as an approximation of W^ϵ given by (2.2). To motivate the definition let us first make some comments. According to the stationary phase method, the main contribution in $W^\epsilon(\mathbf{x},\mathbf{k})$ will come from the stationary points of F. That is, points $\boldsymbol{\sigma}_0$ for which $\nabla_{\boldsymbol{\sigma}} F(\mathbf{x},\mathbf{k};\boldsymbol{\sigma}_0) = 0$. We have that

$$\nabla_{\boldsymbol{\sigma}} F = \nabla_{\mathbf{x}} S(\mathbf{x}+\boldsymbol{\sigma}) + \nabla_{\mathbf{x}} S(\mathbf{x}-\boldsymbol{\sigma}) - 2\mathbf{k}.$$

The stationary points of F always come in pairs $\pm\boldsymbol{\sigma}_0(\mathbf{x},\mathbf{k})$, and as $\mathbf{k}\to\nabla S(\mathbf{x})$, there always exist stationary points with $\boldsymbol{\sigma}_0(\mathbf{x},\mathbf{k})\to 0$. In particular, when $\mathbf{k}=\nabla S(\mathbf{x})$, then $\boldsymbol{\sigma}_0 = 0$ is always a degenerate critical point. This, of course, is due to the fact that $F(\mathbf{x},\mathbf{k};\boldsymbol{\sigma})$ is an odd function of $\boldsymbol{\sigma}$, that is, $F(\mathbf{x},\mathbf{k};\boldsymbol{\sigma}) = -F(\mathbf{x},\mathbf{k};-\boldsymbol{\sigma})$, and therefore all even derivatives at $\boldsymbol{\sigma}=0$ are equal to zero (for any $\mathbf{x}$, $\mathbf{k}$).

From now on we call the set of points $\Lambda = \{(\mathbf{x},\nabla S(\mathbf{x}))\}$, that is, the graph of $\mathbf{k}=\nabla S(\mathbf{x})$, the Lagrangian manifold-by noting that on Λ the 2-form $d\mathbf{x}\wedge d\mathbf{k}$ vanishes identically.

Although the structure of the critical set of F can be quite complicated (even in the case $N=1$ [3, 8]), we expect that, modulo highly oscillatory terms that tend weakly to zero as $\epsilon\to 0$, the main contribution to the asymptotics of W^ϵ comes from points close to the Lagrangian manifold, that is, $\mathbf{k}\approx\nabla S(\mathbf{x})$.

This is in agreement with the fact that as $\epsilon\to 0$ then $W^\epsilon\to W^0 = A^2(\mathbf{x})\delta(\mathbf{k}-\nabla S(\mathbf{x}))$, a Dirac mass concentrated on the Lagrangian manifold. Also, let us consider the projection identity (1.2):

$$|\psi^\epsilon(\mathbf{x})|^2 = \int_{\mathbf{R}^N} W^\epsilon(\mathbf{x},\mathbf{k})\,d\mathbf{k} = \frac{1}{(\epsilon\pi)^N}\int_{\mathbf{R}^N}\int_{\mathbf{R}^N} D(\mathbf{x},\boldsymbol{\sigma})e^{i\frac{F(\mathbf{x},\mathbf{k};\boldsymbol{\sigma})}{\epsilon}}\,d\boldsymbol{\sigma}\,d\mathbf{k}. \tag{2.5}$$

The main contribution in the calculation of the last integral (over $\mathbf{R}^N_{\boldsymbol{\sigma}}\times\mathbf{R}^N_{\mathbf{k}}$) will come from the points $(\boldsymbol{\sigma},\mathbf{k})$ at which $\nabla_{\boldsymbol{\sigma},\mathbf{k}}F=0$. These are easily is seen to be $\boldsymbol{\sigma}=0$, $\mathbf{k}=\nabla S(\mathbf{x})$. At each $\mathbf{x}$ these critical points are always non-degenerate, since the Hessian matrix $H(F)$ is nonsingular. Indeed for any $i,j=1,2,\ldots,N$, we easily compute that at $(\mathbf{x},\boldsymbol{\sigma}=0,\mathbf{k}=\nabla S(\mathbf{x}))$, $F_{k_i,k_j}=0$, $F_{\sigma_i,\sigma_j}=0$, $F_{\sigma_i,k_j}=-2\delta_{ij}$, and $|\det H(F)|=2^{2N}$. Thus in the calculation of $|\psi^\epsilon(\mathbf{x})|$, the main contribution also comes from the points near the Lagrangian manifold.

Once we restrict attention to points $\mathbf{k}\approx\nabla S(\mathbf{x})$, it is natural to approximate D and F by their Taylor expansion about $\boldsymbol{\sigma}=0$:

$$D(\mathbf{x};\boldsymbol{\sigma}) = A^2(\mathbf{x}) + O(|\boldsymbol{\sigma}|^2),$$

and

$$F(\mathbf{x},\mathbf{k};\boldsymbol{\sigma}) = 2(\nabla S(\mathbf{x})-\mathbf{k})\cdot\sigma + \frac{1}{3}\sum_{i,j,k=1}^{N}\frac{\partial^3 S(\mathbf{x})}{\partial x_i\,\partial x_j\,\partial x_k}\sigma_i\sigma_j\sigma_k + O(|\boldsymbol{\sigma}|^5). \tag{2.6}$$

If we keep only the linear term in the right hand side of (2.6) and replace $D(\mathbf{x};\boldsymbol{\sigma})$ by $A^2(\mathbf{x})$ in (2.2) we would "approximate" $W^\epsilon(\mathbf{x},\mathbf{k})$ by

$$\frac{A^2(\mathbf{x})}{(\epsilon\pi)^N}\int_{\mathbf{R}^N} e^{-\frac{i}{\epsilon}2(\mathbf{k}-\nabla S(\mathbf{x}))\cdot\boldsymbol{\sigma}}\, d\boldsymbol{\sigma} = A^2(\mathbf{x})\delta(\mathbf{k}-\nabla S(\mathbf{x})) = W^0(\mathbf{x},\mathbf{k}),$$

that is, the limit Wigner Dirac mass.

To obtain then a nontrivial approximation, we keep the cubic terms in (2.6), and define the semi-classical Wigner function $\widetilde{W}^\epsilon$, as

$$\widetilde{W}^\epsilon(\mathbf{x},\mathbf{k}) := \frac{A^2(\mathbf{x})}{(\epsilon\pi)^N}\int_{\mathbf{R}^N} e^{\frac{i}{\epsilon}[-2(\mathbf{k}-\nabla S(\mathbf{x}))\cdot\boldsymbol{\sigma}+\frac{1}{3}\sum_{i,j,k=1}^N \frac{\partial^3 S(\mathbf{x})}{\partial x_i\partial x_j\partial x_k}\sigma_i\sigma_j\sigma_k]}\, d\boldsymbol{\sigma}. \tag{2.7}$$

It should be noted that the term "semi-classical" is not a standard term. For instance, in the Physics literature, the scaled Wigner function W^ϵ is sometimes called semi-classical Wigner function whereas in Mathematics literature the term "semi-classical" sometimes refers to the limit Wigner function W^0.

It is convenient at this point to introduce some notation that we will use throughout the rest of this work. If we denote by $c_{ijk} = S_{x_i,x_j,x_k}$ the derivatives of S, and use the summation convention, the cubic form in (2.6) or (2.7), takes the form $g(\boldsymbol{\sigma}) = \frac{1}{3}c_{ijk}\sigma_i\sigma_j\sigma_k$. Since the tensor c_{ijk} is symmetric (the value of c_{ijk} is independent of the order of the indices), we may define N symmetric matrices $\mathbf{C}_k$, $k = 1,2\ldots N$ with elements c_{ijk}. The cubic form then, can also be written in vector form as $g(\boldsymbol{\sigma}) = \frac{1}{3}\boldsymbol{\sigma}^T\mathbf{C}_k\boldsymbol{\sigma}\sigma_k$.

Concerning the phase integral appearing in (2.7), we set

$$P_N(\mathbf{z},\mathbf{C}_k) := \frac{1}{(2\pi)^N}\int_{\mathbf{R}^N} e^{i[\mathbf{z}^T\boldsymbol{\sigma}+\frac{1}{3}\boldsymbol{\sigma}^T\mathbf{C}_k\boldsymbol{\sigma}\sigma_k]}\, d\boldsymbol{\sigma}. \tag{2.8}$$

We also use the notation $P_N(\mathbf{z},c_{ijk})$, or simply $P_N(\mathbf{z})$ when there is no confusion as to what the coefficients c_{ijk} are.

Notice that for $N = 1$, if $c \neq 0$ is a constant, than $P_1(z,c) = |c|^{-1/3}\mathrm{Ai}(zc^{-1/3})$, where $Ai(z)$, is the standard Airy function. In particular $P_1(z,1) = \mathrm{Ai}(z)$. Also, if $\mathbf{C}_k \equiv 0$, it follows from (2.8) that

$$P_N(\mathbf{z},0) = \delta(\mathbf{z}). \tag{2.9}$$

Using this notation, the semi-classical Wigner function (2.7) is written as

$$\widetilde{W}^\epsilon(\mathbf{x},\mathbf{k}) = \left(\frac{2}{\epsilon^{\frac{2}{3}}}\right)^N A^2(\mathbf{x})P_N\left(-\frac{2}{\epsilon^{\frac{2}{3}}}(\mathbf{k}-\nabla S(\mathbf{x})),\ \frac{\partial^3 S(\mathbf{x})}{\partial x_i\partial x_j\partial x_k}\right). \tag{2.10}$$

In the case $N = 1$ this takes the form

$$\widetilde{W}^\epsilon(x,k) = \frac{2}{\epsilon^{\frac{2}{3}}}\frac{A^2(x)}{|\, S'''(x)\,|^{\frac{1}{3}}}Ai\Big(-\frac{2}{\epsilon^{\frac{2}{3}}}\frac{k-S'(x)}{(S'''(x))^{\frac{1}{3}}}\Big), \tag{2.11}$$

and we recover the Airy asymptotics of Filippas & Makrakis [8].

A natural question concerned with the definition (2.7) is whether the integral P_N is well defined. This is addressed in appendix A; see Proposition A.1. We show there that P_N is well defined as a distribution whereas if

$$\sum_{i=1}^{N} (\boldsymbol{\sigma}^T \mathbf{C}_i \boldsymbol{\sigma})^2 \neq 0, \qquad \forall \boldsymbol{\sigma} \in \mathbf{R}^N \setminus \{0\}, \tag{2.12}$$

then P_N is a smooth function.

Failure of condition (2.12) is equivalent to the fact that there exists a unit vector $\boldsymbol{v} \in \mathbf{R}^N$ such that $\boldsymbol{v}^T \mathbf{C}_i \boldsymbol{v} = 0$ for all $i = 1, \dots N$. Since $\mathbf{C}_i$ is the second derivative matrix of $\frac{\partial S(\mathbf{x})}{\partial x_i}$ this is equivalent to the fact that for all i's the surfaces $k_i = \frac{\partial S(\mathbf{x})}{\partial x_i}$ have zero normal curvature in the direction of $\boldsymbol{v} \in \mathbf{R}^N_{\mathbf{x}}$. Hence we call a point $\mathbf{x}_0$ on the Lagrangian manifold $\mathbf{k} = \nabla S(\mathbf{x})$ a *degenerate point* if there exists a vector $\boldsymbol{v} \neq 0$ such that all i-components, (that is, all surfaces $k_i = \frac{\partial S(\mathbf{x})}{\partial x_i}$, $i = 1, \dots N$) have zero normal curvature at the point $\mathbf{x}_0$ in direction of $\boldsymbol{v}$.

Since the semi-classical Wigner function is defined via the phase integral P_N (see (2.10)), any property of P_N yields the corresponding property for $\widetilde{W}^\epsilon$. For instance, from (A 6) we have that

$$\widetilde{W}^\epsilon(\mathbf{x}, \mathbf{k}) \to A^2(\mathbf{x}) \delta(\mathbf{k} - \nabla S(\mathbf{x})), \quad \text{as} \quad \epsilon \to 0, \tag{2.13}$$

whereas, if $\mathbf{x}$ is a non-degenerate point, it follows from (A 4) that

$$\int_{\mathbf{R}^N} \widetilde{W}^\epsilon(\mathbf{x}, \mathbf{k})\, d\mathbf{k} = A^2(\mathbf{x}). \tag{2.14}$$

Near non-degenerate points the semi-classical Wigner is a smooth function that approximates the scaled Wigner W^ϵ, as ϵ tends to zero. At degenerate points either the semi-classical Wigner fails to approximate W^ϵ, or else W^ϵ itself is a distribution.

The reason that at a degenerate point the semi-classical Wigner fails to approximate the scaled Wigner, is of course, due to the fact that by truncating the Taylor expansion of F (cf. (2.6)) at the cubic order terms and discarding terms of higher order, we do not always obtain nontrivial approximations of the phase F. This is easily seen in the one dimensional case: When $N = 1$ then $C_1 = S'''(x)$ and failure of condition (2.12) is equivalent to $S'''(x) = 0$. At such a point, keeping the cubic term in (2.6) is of no use and the semi-classical Wigner coincides with the limit Wigner (a distribution). To obtain a nontrivial approximation at a degenerate point one ought to keep more terms in (2.6).

Remark 2.1 To arrive at the definition of $\widetilde{W}^\epsilon$, we (i) threw away oscillatory terms that tend to zero as $\epsilon \to 0$; these terms originate from nonzero ($\boldsymbol{\sigma}_0 \neq 0$) stationary points of F and (ii) we replaced F by its third order Taylor expansion about $\boldsymbol{\sigma} = 0$. Thus, with the exception of highly oscillatory terms due to (i), $\widetilde{W}^\epsilon$ approximates W^ϵ locally, near the Lagrangian manifold. The question of a uniform approximation of W^ϵ, under the generality we consider it here, seems to be an impossible task. We note however that under the assumption that the Lagrangian manifold is globally convex, the oscillatory terms of (i) are absent. In addition, the phase F in (2.4) can be identified with a suitable symplectic area – without using Taylor expansion. One then obtains analytic expressions for the stationary phase approximation of W^ϵ, in terms of suitable geometric quantities, which are valid not only near the manifold but also in larger regions in phase space. We refer to [3] for the case $N = 1$ and [18] for $N = 2$. Finally, it is clear that if $A = constant$ and $S(\mathbf{x})$ is a cubic polynomial in $\mathbf{x}$ then $\widetilde{W}^\epsilon \equiv W^\epsilon$.

Remark 2.2 We note that by keeping two terms in the Taylor expansion (2.6), we have an object ($\widetilde{W}^\epsilon$) that is "richer" than the limit Wigner W^0, in the sense that (i) is ϵ-dependent and (ii) we can always send ϵ to zero to recover W^0; see (2.13). In fact, away from degenerate points, $\widetilde{W}^\epsilon$ provides the proper (P_N) regularization of the limit Wigner Dirac mass. At degenerate points, W^ϵ has a structure more rich than $\widetilde{W}^\epsilon$ can capture. As a result, $\widetilde{W}^\epsilon$ is not a good approximant of W^ϵ in the neighborhood of these points, and is unable to provide the proper regularization of W^0.

3 Phase space dynamics

Our analysis in this section is restricted to the single phase case, that is, in the time interval $0 \leqslant t \leqslant T$ with $T < t_c$, where t_c is the first time a caustic appears. If $\psi_\epsilon(\mathbf{x}, t) = A(\mathbf{x}, t)e^{\frac{i}{\epsilon}S(\mathbf{x},t)}$, is the WKB solution of (1.3), (1.4), the semi-classical Wigner $\widetilde{W}^\epsilon$ of ψ_ϵ at any time $t \in [0, T]$ is given by (2.10). We recall that our objective is to obtain a direct evolution law for $\widetilde{W}^\epsilon$, without having to go through the WKB solution. As a first step, we need to know how the quantities $\mathbf{k} - \nabla S(\mathbf{x}, t)$ and $S_{x_i x_j x_k}(\mathbf{x}, t)$, that enter in the argument of P_N, evolve with time.

Let us first introduce some notation and recall some basic facts. We denote by $(\mathbf{x}, \mathbf{k})$ a point in phase space and we use the special notation $(\mathbf{q}, \mathbf{p})$ for points of phase space at time $t = 0$. The Hamiltonian flow moves a point $(\mathbf{q}, \mathbf{p})$ (at $t = 0$) to the point $(\mathbf{x}, \mathbf{k})$ (at $t > 0$), along the bicharacteristics given by $\mathbf{x} = \hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t)$ and $\mathbf{k} = \hat{\mathbf{k}}(\mathbf{q}, \mathbf{p}, t)$.

For the inverse bicharacteristics we use the notation $\mathbf{q} = \hat{\mathbf{q}}(\mathbf{x}, \mathbf{k}, t)$ and $\mathbf{p} = \hat{\mathbf{p}}(\mathbf{x}, \mathbf{k}, t)$. We use the "bar" notation for the rays $\overline{\mathbf{x}}(\mathbf{q}, t) = \hat{\mathbf{x}}(\mathbf{q}, \nabla S_0(\mathbf{q}), t)$ and similarly for the inverse rays $\overline{\mathbf{q}}(\mathbf{x}, t) = \hat{\mathbf{q}}(\mathbf{x}, \nabla S(\mathbf{x}), t)$.

We recall that the Hamiltonian flow, for the problem at hand, is given by the ODE system

$$\frac{d}{dt}\hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t) = \hat{\mathbf{k}}(\mathbf{q}, \mathbf{p}, t), \qquad \frac{d}{dt}\hat{\mathbf{k}}(\mathbf{q}, \mathbf{p}, t) = -\nabla_{\mathbf{x}} V(\hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t)). \tag{3.1}$$

The initial Lagrangian manifold Λ_0 associated with $\psi_\epsilon(\mathbf{q}) = A_0(\mathbf{q})e^{\frac{i}{\epsilon}S_0(\mathbf{q})}$ is defined by $\Lambda_0 = \{(\mathbf{q}, \mathbf{p}) : \mathbf{p} = \nabla S_0(\mathbf{q})\}$. This is a graph (from $\mathbf{R}^N_{\mathbf{q}}$ to $\mathbf{R}^N_{\mathbf{p}}$) in the sense that to each $\mathbf{q}$ there corresponds a unique $\mathbf{p}$. At time t the Hamiltonian flow moves Λ_0 to $\Lambda_t = \{(\mathbf{x}, \mathbf{k}) : \hat{\mathbf{p}}(\mathbf{x}, \mathbf{k}, t) = \nabla S_0(\hat{\mathbf{q}}(\mathbf{x}, \mathbf{k}, t))\}$. Under our assumption that we are in the single phase case, Λ_t remains a graph and is alternatively given by $\Lambda_t = \{(\mathbf{x}, \mathbf{k}) : \mathbf{k} = \nabla_{\mathbf{x}} S(\mathbf{x}, t)\}$, where $S(\mathbf{x}, t)$ is the (single valued) solution of the eikonal equation (1.6).

We are now in position to start our calculations. Let $\mathbf{a}(t) = \mathbf{a}(\mathbf{q}, \mathbf{p}, t) = \hat{\mathbf{k}}(\mathbf{q}, \mathbf{p}, t) - \nabla_{\mathbf{x}} S(\hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t), t)$ with $\mathbf{a}(0) = \mathbf{p} - \nabla S_0(\mathbf{q})$. We will derive and solve an ODE for $\mathbf{a}(t)$ along the bicharacteristic $(\mathbf{q}, \mathbf{p}) \to (\mathbf{x}, \mathbf{k})$, for $(\mathbf{q}, \mathbf{p})$ close to Λ_0. For simplicity we will sometimes use the notation $\hat{\mathbf{x}}$ and $\hat{\mathbf{k}}$ in the place of $\hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t)$ and $\hat{\mathbf{k}}(\mathbf{q}, \mathbf{p}, t)$ respectively.

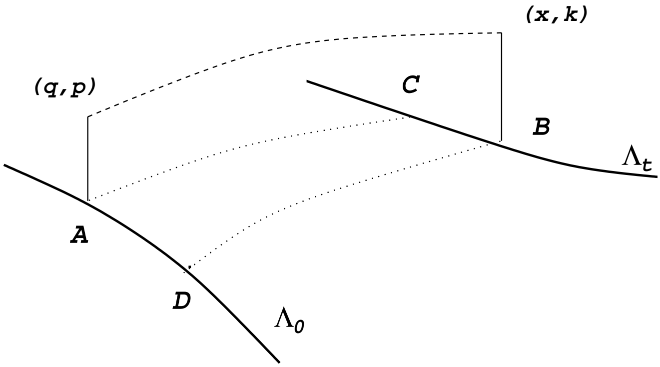

FIGURE 1. Evolution of $\mathbf{k} - \nabla S(\mathbf{x}, t)$. The point $(\mathbf{q}, \mathbf{p})$ moves to $(\mathbf{x}, \mathbf{k}) = (\hat{\mathbf{x}}(t; \mathbf{q}, \mathbf{p}),\ \hat{\mathbf{k}}(t; \mathbf{q}, \mathbf{p}))$. In the figure are also shown the points $A = (\mathbf{q}, \nabla S_0(\mathbf{q}))$, $B = (\mathbf{x}, \nabla S(\mathbf{x}, t))$, $C = (\overline{\mathbf{x}}(\mathbf{q}, t),\ \nabla S(\overline{\mathbf{x}}(\mathbf{q}, t), t)$ and $D = (\overline{\mathbf{q}}(\mathbf{x}, t),\ \nabla S_0(\overline{\mathbf{q}}(\mathbf{x}, t))$.

Differentiating along the bicharacteristic $(\mathbf{q}, \mathbf{p}) \to (\mathbf{x}, \mathbf{k})$ we get

$$\begin{aligned} \frac{d}{dt}\mathbf{a}(t) &= \frac{d}{dt}\hat{\mathbf{k}}(\mathbf{q}, \mathbf{p}, t) - \frac{d}{dt}\nabla_{\mathbf{x}} S(\hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t), t) \\ &= -\nabla V(\hat{\mathbf{x}}) - (\partial_t + \hat{\mathbf{k}} \cdot \nabla)\nabla_{\mathbf{x}} S(\hat{\mathbf{x}}, t) \\ &= -\nabla V(\hat{\mathbf{x}}) - \nabla_{\mathbf{x}} S_t(\hat{\mathbf{x}}, t) - \mathbf{B}(t) \cdot \hat{\mathbf{k}}, \end{aligned} \tag{3.2}$$

where $\mathbf{B}(t)$ is the symmetric $N \times N$ matrix with elements $b_{ij}(t) = S_{x_i x_j}(\hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t), t)$. Taking the gradient of the eikonal equation (1.6), we find for $S(\hat{\mathbf{x}}, t) = S(\hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t), t)$

$$\nabla_{\mathbf{x}} S_t(\hat{\mathbf{x}}, t) = -\mathbf{B}(t) \cdot \nabla_{\mathbf{x}} S(\hat{\mathbf{x}}, t) - \nabla V(\hat{\mathbf{x}}). \tag{3.3}$$

From (3.2) and (3.3) we see that

$$\frac{d}{dt}\mathbf{a}(t) = -\mathbf{B}(t)\mathbf{a}(t). \tag{3.4}$$

If we denote by $\Phi(t) = \{\phi_{ij}(t)\}_{i,j=1,\ldots,N}$ the fundamental solution of (3.4) with $\Phi(0) = \mathbf{I}_N$, then

$$\mathbf{a}(t) = \Phi(t)\mathbf{a}(0). \tag{3.5}$$

For future use we need to estimate $\det \Phi(t)$. To this end we start with a well known formula about the Jacobian of a ray $\overline{\mathbf{x}}(\mathbf{q}, t)$ (cf. *AC* in Figure 1); see (3.6) below. We present a proof of this formula for completeness. Let $\overline{\mathbf{x}}(\mathbf{q}, t) = (\overline{x}_i(\mathbf{q}, t), \ldots, \overline{x}_N(\mathbf{q}, t))$. Then

$$\frac{\partial \overline{x}_i(\mathbf{q}, t)}{\partial t} = \hat{k}_i(\mathbf{q}, \nabla S_0(\mathbf{q}), t) = S_{x_i}(\overline{\mathbf{x}}(\mathbf{q}, t), t).$$

Taking the $\frac{\partial}{\partial q_j}$ derivative of this we find

$$\frac{\partial}{\partial t}\left(\frac{\partial \overline{x}_i}{\partial q_j}\right) = \sum_{k=1}^{N} S_{x_i x_k} \frac{\partial \overline{x}_k}{\partial q_j}.$$

Setting $\overline{\mathbf{X}}(t)$ for the matrix with elements $\overline{X}_{ij} = \frac{\partial \overline{x}_i}{\partial q_j}$ and $\overline{\mathbf{B}}(t)$ for the symmetric matrix with elements $S_{x_i x_j}(\overline{\mathbf{x}}(\mathbf{q}, t), t)$, we can rewrite the above system as

$$\frac{d}{dt}\overline{\mathbf{X}}(t) = \overline{\mathbf{B}}(t)\overline{\mathbf{X}}(t).$$

But $J(\mathbf{q}, t) = \det \overline{\mathbf{X}}(t)$ is the Jacobian of the ray $\overline{\mathbf{x}}(\mathbf{q}, t)$, and by standard ODE theory (cf. Coddington & Levinson [6]) we have that

$$J(\mathbf{q}, t) = \exp\left\{ \int_0^t \operatorname{tr} \overline{\mathbf{B}}(\tau) d\tau \right\} = \exp\left\{ \int_0^t \Delta_{\mathbf{x}} S(\overline{\mathbf{x}}(\mathbf{q}, \tau), \tau) d\tau \right\}. \tag{3.6}$$

Since $\Phi(t)$ solves (3.4) we also have

$$\det \Phi(t) = \exp\left\{ -\int_0^t \operatorname{tr} \mathbf{B}(\tau) d\tau \right\} = \exp\left\{ -\int_0^t \Delta_{\mathbf{x}} S(\hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, \tau), \tau) d\tau \right\}. \tag{3.7}$$

To relate $J(\mathbf{q}, t)$ and $\det \Phi(t)$ we will use the following approximation:

$$S_{x_i x_i}(\hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t), t) = S_{x_i x_i}(\hat{\mathbf{x}}(\mathbf{q}, S_0(\mathbf{q}), t), t) + O(t|\mathbf{a}(0)|). \tag{3.8}$$

If we accept this we easily arrive at

$$\det \Phi(t) = J^{-1}(\mathbf{q}, t)(1 + O(t^2|\mathbf{a}(0)|)). \tag{3.9}$$

It remains to prove estimate (3.8). For easier comparison it is more convenient at this point to use the bicharacteristic notation $\hat{\mathbf{x}}(\mathbf{q}, S_0(\mathbf{q}), t)$ instead of the ray notation. For $\mathbf{p}$ close to $S_0(\mathbf{q})$ we have that

$$S_{x_i x_i}(\hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t), t) - S_{x_i x_i}(\hat{\mathbf{x}}(\mathbf{q}, S_0(\mathbf{q}), t), t) = O(|\hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t) - \hat{\mathbf{x}}(\mathbf{q}, S_0(\mathbf{q}), t)|). \tag{3.10}$$

On the other hand,

$$|\hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t) - \hat{\mathbf{x}}(\mathbf{q}, S_0(\mathbf{q}), t)| = O(|\nabla_{\mathbf{p}} \hat{\mathbf{x}}(\mathbf{q}, S_0(\mathbf{q}), t)||\mathbf{p} - S_0(\mathbf{q})|). \tag{3.11}$$

From the Hamiltonian system it follows that

$$\frac{d}{dt}\nabla_{\mathbf{p}} \hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t) = \nabla_{\mathbf{p}} \hat{\mathbf{k}}(\mathbf{q}, \mathbf{p}, t), \quad \text{with} \quad \nabla_{\mathbf{p}} \hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, 0) = 0, \quad \nabla_{\mathbf{p}} \hat{\mathbf{k}}(\mathbf{q}, \mathbf{p}, 0) = I_N.$$

Hence, for t small we have that $|\nabla_{\mathbf{p}} \hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t)| = O(t)$. Clearly, this estimate is also valid in the (compact) time interval $0 \leqslant t \leqslant T$. Recalling that $\mathbf{a}(0) = \mathbf{p} - \nabla S_0(\mathbf{q})$ we then have from (3.11) that

$$|\hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t) - \hat{\mathbf{x}}(\mathbf{q}, S_0(\mathbf{q}), t)| = O(t|\mathbf{a}(0)|),$$

and (3.8) follows from (3.11).

We next derive ODEs for the evolution of the third derivatives. Let

$$c_{ijk}(t) = S_{x_i x_j x_k}(\hat{\mathbf{x}}(\mathbf{q}, \mathbf{p}, t), t),$$

with $c_{ijk}(0) = S_{0,q_iq_jq_k}(\mathbf{q})$. Differentiating along the bicharacteristic $(\mathbf{q},\mathbf{p}) \to (\mathbf{x},\mathbf{k})$ we get for $S = S(\hat{\mathbf{x}}(\mathbf{q},\mathbf{p},t),t)$

$$\frac{d}{dt}c_{ijk}(t) = (\partial_t + \hat{\mathbf{k}}\cdot\nabla_{\mathbf{x}})S_{x_ix_jx_k} = S_{x_ix_jx_k,t} + \hat{k}_l S_{x_ix_jx_kx_l}. \tag{3.12}$$

Taking the $\frac{\partial^3}{\partial x_i\partial x_j\partial x_k}$– derivative of the eikonal (1.6) we get for $S = S(\hat{\mathbf{x}},t) = S(\hat{\mathbf{x}}(\mathbf{q},\mathbf{p},t),t)$

$$S_{x_ix_jx_k,t} = -(S_{x_lx_i}S_{x_lx_jx_k} + S_{x_lx_j}S_{x_lx_ix_k} + S_{x_lx_k}S_{x_lx_ix_j}) - S_{x_l}S_{x_lx_ix_jx_k} - V_{x_ix_jx_k},$$

where $V_{x_ix_jx_k} = V_{x_ix_jx_k}(\hat{\mathbf{x}}(\mathbf{q},\mathbf{p},t),t)$. Replacing $S_{x_ix_jx_k,t}$ in (3.12) and using our notation we end up with

$$\begin{aligned}\frac{d}{dt}c_{ijk} &= -(b_{li}c_{ljk} + b_{lj}c_{lik} + b_{lk}c_{lij}) + \mathbf{a}(t)\cdot\nabla_{\mathbf{x}}S_{x_ix_jx_k} - V_{x_ix_jx_k}\\ &= -(b_{li}c_{ljk} + b_{lj}c_{lik} + b_{lk}c_{lij}) - V_{x_ix_jx_k} + O(t|\mathbf{a}(t)|).\end{aligned} \tag{3.13}$$

Let us denote by $\mathbf{C}_k(t)$ and $\mathbf{V}_k(t)$, $k = 1,2,\dots,N$ the symmetric $N\times N$ matrices with elements $c_{ijk}(t)$ and $V_{x_ix_jx_k}$ $i,j = 1,2,\dots,N$ respectively. We then write the equation (3.13) in matrix form as $(k,l = 1,2,\dots N)$

$$\frac{d}{dt}\mathbf{C}_k(t) = -\mathbf{B}(t)\mathbf{C}_k(t) - \mathbf{C}_k(t)\mathbf{B}(t) - b_{kl}(t)\mathbf{C}_l(t) - \mathbf{V}_k(t) + O(t|\mathbf{a}(0)|). \tag{3.14}$$

By the results of Appendix B, the solution of (3.14) is given by $(k,l = 1,2,\dots N)$

$$\mathbf{C}_k(t) = \Phi(t)\left[\mathbf{C}_l(0) + \mathbf{U}_l(t) + O(t|\mathbf{a}(0)|)\right]\Phi^T(t)\phi_{kl}(t), \tag{3.15}$$

where $\mathbf{U}_k(t)$ solves

$$\phi_{kl}\frac{d}{dt}\mathbf{U}_l = -\Phi^{-1}\mathbf{V}_k\Phi^{-T}, \qquad \mathbf{U}_k(0) = 0, \qquad k,l = 1,2,\dots,N. \tag{3.16}$$

We may write (3.15) as

$$\mathbf{C}_k(t) + O(t|\mathbf{a}(0)|) = \Phi(t)\left[\mathbf{C}_l(0) + \mathbf{U}_l(t)\right]\Phi^T(t)\phi_{kl}(t). \tag{3.17}$$

Remark 3.1 We recall that the calculation of $\mathbf{B}(t)$, $\Phi(t)$, $\mathbf{C}_k(t)$ and $\mathbf{U}(t)$ have been done along the bicharacteristic $(\mathbf{q},\mathbf{p}) \to (\mathbf{x},\mathbf{k})$. To keep track of the bicharacteristic, when needed, we will use the notation $\mathbf{B}(\mathbf{q},\mathbf{p},t)$, $\Phi(\mathbf{q},\mathbf{p},t)$, etc.

Remark 3.2 Let us use the "bar" notation when the previous quantities are calculated along rays. For instance, $\overline{c}_{ijk}(t) = S_{x_ix_jx_k}(\hat{\mathbf{x}}(\mathbf{q},\nabla S_0(\mathbf{q}),t),t) = S_{x_ix_jx_k}(\overline{\mathbf{x}}(\mathbf{q},t),t)$, and similarly for $\overline{V}_{x_ix_jx_k}$, $\overline{\mathbf{C}}_k$, $\overline{\mathbf{V}}_k$ etc. Then, the equation for $\overline{\mathbf{C}}_k$ is the same as (3.14) without the error terms and in particular, it follows from (3.15) that

$$\overline{\mathbf{C}}_k(t) = \overline{\Phi}(t)\left[\overline{\mathbf{C}}_l(0) + \overline{\mathbf{U}}_l(t))\right]\overline{\Phi}^T(t)\overline{\phi}_{kl}(t); \tag{3.18}$$

here the matrix $\overline{\Phi}(t)$ is the fundamental solution of the equation

$$\frac{d}{dt}\mathbf{a}(t) = -\overline{\mathbf{B}}(t)\mathbf{a}(t). \tag{3.19}$$

Also, working as above, we can derive the ODE satisfied by $\overline{\mathbf{B}}(t)$:

$$\frac{d}{dt}\overline{\mathbf{B}}(t) = -\overline{\mathbf{B}}^2(t) - \overline{\mathbf{V}}(t), \tag{3.20}$$

where $\overline{\mathbf{V}} = \overline{V}_{x_i x_j}$ is the second derivative matrix of V.

4 Evolution of the limit Wigner function

Here we will use the phase-space formulas derived above, to solve the Liouville equation for the limit Wigner function (cf. (1.14)).

In the sequel we will write $\hat{\mathbf{q}}$ instead of $\hat{\mathbf{q}}(\mathbf{x},\mathbf{k},t)$, and $\hat{\mathbf{p}}$ instead of $\hat{\mathbf{p}}(\mathbf{x},\mathbf{k},t)$. We also set $\hat{\mathbf{a}}(t) := \mathbf{k}(\hat{\mathbf{q}},\hat{\mathbf{p}},t) - \nabla_{\mathbf{x}}S(\mathbf{x}(\hat{\mathbf{q}},\hat{\mathbf{p}},t),t) = \mathbf{k} - \nabla_{\mathbf{x}}S(\mathbf{x},t)$; in particular $\hat{\mathbf{a}}(0) = \hat{\mathbf{p}} - \nabla S_0(\hat{\mathbf{q}})$. We note that $\hat{\mathbf{a}}$ and $\mathbf{a}$ (defined in section 3) represent the same quantity; $\hat{\mathbf{a}}$ is a function of $(\mathbf{x},\mathbf{k},t)$ and $\mathbf{a}$ is a function of $(\mathbf{q},\mathbf{p},t)$. Similarly, we set $\hat{\Phi}(t) := \Phi(\hat{\mathbf{q}},\hat{\mathbf{p}},t)$.

Solving (1.14) by characteristics we get

$$\begin{aligned} f^0(\mathbf{x},\mathbf{k},t) &= f_0^0(\hat{\mathbf{q}}(t;\mathbf{x},\mathbf{k}),\hat{\mathbf{p}}(t;\mathbf{x},\mathbf{k})) \\ &= A_0^2(\hat{\mathbf{q}})\delta(\hat{\mathbf{p}} - \nabla S_0(\hat{\mathbf{q}})) \\ &= A_0^2(\hat{\mathbf{q}})\delta(\hat{\mathbf{a}}(0)). \end{aligned} \tag{4.1}$$

When restricted in the single phase region $[0,T]$, $T < t_c$, we use (3.5), (3.9) and (1.9) to obtain

$$\begin{aligned} A_0^2(\hat{\mathbf{q}})\delta(\hat{\mathbf{a}}(0)) &= A_0^2(\hat{\mathbf{q}})\delta(\hat{\Phi}^{-1}(t)\hat{\mathbf{a}}(t)) \\ &= A_0^2(\hat{\mathbf{q}})|\det\hat{\Phi}(t)|\delta(\hat{\mathbf{a}}(t)) \\ &= \frac{A_0^2(\hat{\mathbf{q}})}{J(\hat{\mathbf{q}},t)}\,\delta(\hat{\mathbf{a}}(t)) \\ &= A^2(\mathbf{x},t)\delta(\mathbf{k} - \nabla S_{\mathbf{x}}(\mathbf{x},t)), \end{aligned} \tag{4.2}$$

recovering the well known fact that f^0 coincides with the limit Wigner of the WKB solution.

Formula (4.1) is valid for any $t \geqslant 0$ even in the multiphase case. As noted in [8], at a caustic point $(\mathbf{x}_c,t_c)$ the quantity $\hat{\mathbf{q}}(\mathbf{x}_c,\mathbf{k},t_c) - \nabla S_0(\hat{\mathbf{q}}(\mathbf{x}_c,\mathbf{k},t_c))$ ceases to have simple roots with respect to $\mathbf{k}$ and the composition of the Dirac mass with $\hat{\mathbf{q}}(\mathbf{x}_c,\mathbf{k},t) - \nabla S_0(\hat{\mathbf{q}}(\mathbf{x}_c,\mathbf{k},t_c))$ is not well defined as a measure in $\mathbf{R}^N_{\mathbf{k}}$, not even as a distribution in $\mathbf{R}^N_{\mathbf{k}}$ [15, p. 547]. This means that at the point $(\mathbf{x}_c,t_c)$, the integral $\int_{\mathbf{R}^N_{\mathbf{k}}} f^0(\mathbf{x}_c,\mathbf{k},t_c)d\mathbf{k}$ is meaningless and consequently, at such a point, the projection formula (1.17) is inapplicable.

To motivate the definition of $\widetilde{f}^\epsilon$ that will be given in the next section, let us make some observations about f^0. Suppose in the Taylor expansion (2.6) we keep only the linear term, in which case we "approximate" W^ϵ by the limit Wigner W^0. Thus – through comparison with the scaled Wigner of the WKB solution – we would end up with the well known formula

$$W^0(\mathbf{x},\mathbf{k},t) = A(\mathbf{x},t)\delta(\mathbf{k} - \nabla S_\mathbf{x}(\mathbf{x},t)), \tag{4.3}$$

valid for any $t \geqslant 0$, with $A(\mathbf{x},t)$ and $S(\mathbf{x},t)$ as given by the WKB method. Suppose now that we are unaware of the fact that W^0 solves a Liouville equation in phase space, and we try to finger out what is the evolution law of W^0. By comparing W^0 at time $t > 0$ and at $t = 0$, using relations (3.5), (3.9) and (1.9) one would immediately see that W^0 is transported by the Hamiltonian flow – by the same calculations as in (4.1), (4.2), but done in reverse order. Hence, we can now "define" f^0 by applying the correct evolution law (transport) on W_0^0, that is, $f^0(\mathbf{x},\mathbf{k},t) = W_0^0(\hat{\mathbf{q}}(\mathbf{x},\mathbf{k},t),\hat{\mathbf{p}}(\mathbf{x},\mathbf{k},t))$, and this of course, leads to the correct definition.

In the next section we will define $\widetilde{f}^\epsilon$ by applying a suitable evolution law to $\widetilde{W}_0^\epsilon(\mathbf{q},\mathbf{p})$. The evolution law will follow by comparing $\widetilde{W}^\epsilon$ at time $t > 0$ and $t = 0$, using the phase space evolution formulas of § 3 as well as (1.9) and (A 12).

5 Evolution of the semi-classical Wigner function

Here we will derive the evolution law for the semi-classical Wigner function. We will use the notation of the previous section, that is, $\hat{\mathbf{q}}$, $\hat{\mathbf{p}}$, $\hat{\mathbf{a}}(t)$ and $\hat{\Phi}(t)$. In addition we set $\hat{\mathbf{C}}_k(t) := \mathbf{C}_k(\hat{\mathbf{q}},\hat{\mathbf{p}},t) = \mathbf{C}_k(\hat{\mathbf{q}}(\mathbf{x},\mathbf{k},t),\hat{\mathbf{p}}(\mathbf{x},\mathbf{k},t),t)$ and similarly for $\hat{\mathbf{V}}_k(t)$ and $\hat{\mathbf{U}}_k(t)$.

Our starting point is the semi-classical Wigner of the WKB solution at time t, that is,

$$\widetilde{W}^\epsilon(\mathbf{x},\mathbf{k},t) = \left(2\epsilon^{-\frac{2}{3}}\right)^N A^2(\mathbf{x},t) P_N\left(-2\epsilon^{-\frac{2}{3}}\hat{\mathbf{a}}(t),\hat{\mathbf{C}}_k(t)\right).$$

Recalling (3.5) and (3.15) we have that

$$\hat{\mathbf{a}}(t) = \hat{\Phi}(t)\hat{\mathbf{a}}(0), \qquad \hat{\mathbf{C}}_k(t) = \hat{\Phi}(t)[\hat{\mathbf{C}}_l(0) + \hat{\mathbf{U}}_l(t) + O(t|\hat{\mathbf{a}}(0)|)]\hat{\Phi}^T(t)\hat{\phi}_{kl}(t).$$

Using the change of variables formula (cf. (A 8), (A 9) with $\mathbf{T} = \hat{\Phi}^T(t)$) first and then the convolution formula (cf. (A 12)), with $\eta = -2\epsilon^{-\frac{2}{3}}$, $\mathbf{z} = \mathbf{p}$, $\boldsymbol{\alpha} = \nabla S_0(\hat{\mathbf{q}})$, we get

$$\begin{aligned}
&|\det\hat{\Phi}(t)|\; P_N\left(-2\epsilon^{-\frac{2}{3}}\hat{\mathbf{a}}(t),\ \hat{\mathbf{C}}_k(t)\right)\\
&\quad = P_N\left(-2\epsilon^{-\frac{2}{3}}\hat{\mathbf{a}}(0),\ \hat{\mathbf{C}}_k(0) + \hat{\mathbf{U}}_k(t) + O(t|\hat{\mathbf{a}}(0)|\right)\\
&\quad = P_N\left(-2\epsilon^{-\frac{2}{3}}(\hat{\mathbf{p}} - \nabla S_0(\hat{\mathbf{q}})),\ \hat{\mathbf{C}}_k(0) + \hat{\mathbf{U}}_k(t) + O(t|\hat{\mathbf{a}}(0)|\right)\\
&\quad = P_N\left(-2\epsilon^{-\frac{2}{3}}(\mathbf{p} - \nabla S_0(\hat{\mathbf{q}})),\ \hat{\mathbf{C}}_k(0)\right) *_\mathbf{p} P_N\left(-2\epsilon^{-\frac{2}{3}}\mathbf{p}, \hat{\mathbf{U}}_k(t) + O(t|\hat{\mathbf{a}}(0)|)\right)\Big|_{\mathbf{p}=\hat{\mathbf{p}}}\ \left(2\epsilon^{-\frac{2}{3}}\right)^N.
\end{aligned}$$

From this and (3.9) it follows that

$$
\begin{aligned}
&\widetilde{W}^{\epsilon}(\mathbf{x},\mathbf{k},t)\\
&= \left(2\epsilon^{-\frac{2}{3}}\right)^{N} A^2(\mathbf{x},t)J(\hat{\mathbf{q}},t)(1+O(t^2|\hat{\mathbf{a}}(0)|)) \times\\
&\quad P_N\left(-2\epsilon^{-\frac{2}{3}}(\mathbf{p}-\nabla S_0(\hat{\mathbf{q}})),\ \hat{\mathbf{C}}_k(0)\right) *_{\mathbf{p}} P_N\left(-2\epsilon^{-\frac{2}{3}}\mathbf{p}, \hat{\mathbf{U}}_k(t)+O(t|\hat{\mathbf{a}}(0)|)\right)\Big|_{\mathbf{p}=\hat{\mathbf{p}}} \left(2\epsilon^{-\frac{2}{3}}\right)^{N}\\
&= \left(2\epsilon^{-\frac{2}{3}}\right)^{N} A_0^2(\hat{\mathbf{q}})P_N\left(-2\epsilon^{-\frac{2}{3}}(\mathbf{p}-\nabla S_0(\hat{\mathbf{q}})), \hat{\mathbf{C}}_k(0)\right) *_{\mathbf{p}}\\
&\quad *_{\mathbf{p}} P_N\left(-2\epsilon^{-\frac{2}{3}}\mathbf{p},\ \hat{\mathbf{U}}_k(t)+O(t|\hat{\mathbf{a}}(0)|)\right)\Big|_{\mathbf{p}=\hat{\mathbf{p}}}(1+O(t^2|\hat{\mathbf{a}}(0)|))\left(2\epsilon^{-\frac{2}{3}}\right)^{N}\\
&= \widetilde{W}_0^{\epsilon}(\hat{\mathbf{q}},\mathbf{p}) *_{\mathbf{p}} P_N\left(-2\epsilon^{-\frac{2}{3}}\mathbf{p},\ \hat{\mathbf{U}}_k(t)+O(t|\hat{\mathbf{a}}(0)|)\right)\Big|_{\mathbf{p}=\hat{\mathbf{p}}}(1+O(t^2|\hat{\mathbf{a}}(0)|))\left(2\epsilon^{-\frac{2}{3}}\right)^{N}.
\end{aligned}
$$

Thus, with the exception of the error terms, $\widetilde{W}^{\epsilon}(\mathbf{x},\mathbf{k},t)$ is equal to the convolution of $\widetilde{W}_0^{\epsilon}(\hat{\mathbf{q}},\mathbf{p})$ with a suitable phase integral, the convolution being evaluated at the point $\mathbf{p}=\hat{\mathbf{p}}$.

Omitting the error terms in the last line of the calculations above, we set

$$
\tilde{G}^{\epsilon}(\mathbf{p},\hat{\mathbf{U}}_k(t)) := \left(2\epsilon^{-\frac{2}{3}}\right)^{N} P_N(-2\epsilon^{-\frac{2}{3}}\mathbf{p},\hat{\mathbf{U}}_k(t)). \tag{5.1}
$$

We then define $\widetilde{f}^{\epsilon}(\mathbf{x},\mathbf{k},t)$ as

$$
\widetilde{f}^{\epsilon}(\mathbf{x},\mathbf{k},t) := \tilde{G}^{\epsilon}(\mathbf{p},\hat{\mathbf{U}}_k(t)) *_{\mathbf{p}} \widetilde{W}_0^{\epsilon}(\hat{\mathbf{q}},\mathbf{p})\Big|_{\mathbf{p}=\hat{\mathbf{p}}} = \tilde{G}^{\epsilon}(\mathbf{p},\hat{\mathbf{U}}_k(t)) *_{\mathbf{p}} \widetilde{f}_0^{\epsilon}(\hat{\mathbf{q}},\mathbf{p})\Big|_{\mathbf{p}=\hat{\mathbf{p}}}. \tag{5.2}
$$

We note that $\widetilde{f}^{\epsilon}(\mathbf{x},\mathbf{k},t)$ as defined by (5.2) is in agreement with $\widetilde{W}^{\epsilon}(\mathbf{x},\mathbf{k},t)$, for $|\hat{\mathbf{a}}(0)|$ small (cf. (A 7)), that is, locally near the Lagrangian manifold.

Let us recall that $\hat{\mathbf{U}}_k(t)$ is a known quantity, in the sense that one has first to solve the ODE (3.2) to get $\hat{\Phi}(t)$ and then solve the ODE (3.16) to get $\hat{\mathbf{U}}_k(t)$. Both ODE's are solved along the bicharacteristics $(\hat{\mathbf{q}},\hat{\mathbf{p}}) \to (\mathbf{x},\mathbf{k})$. It follows easily that $\tilde{G}^{\epsilon}$ depends on the potential $V(\mathbf{x})$ and the initial phase S_0, but is independent of the initial amplitude A_0.

The evolved semi-classical Wigner function as defined by (5.2) is in fact a P_N–phase integral, given by

$$
\widetilde{f}^{\epsilon}(\mathbf{x},\mathbf{k},t) = \left(2\epsilon^{-\frac{2}{3}}\right)^{N} A_0^2(\hat{\mathbf{q}})P_N\left(-2\epsilon^{-\frac{2}{3}}\hat{\mathbf{a}}(0),\ \hat{\mathbf{C}}_k(0)+\hat{\mathbf{U}}_k(t)\right), \tag{5.3}
$$

and it follows easily that

$$
\widetilde{f}^{\epsilon}(\mathbf{x},\mathbf{k},t) \to f^0(\mathbf{x},\mathbf{k},t) = A^2(\mathbf{x},t)\delta(\mathbf{k}-\nabla_{\mathbf{x}}S(\mathbf{x},t)), \quad \text{as} \quad \epsilon \to 0.
$$

We also note that $\widetilde{f}^{\epsilon}(\mathbf{x},\mathbf{k},t)$ is a smooth function iff the point $(\mathbf{x},\nabla S(\mathbf{x},t))$ on Λ_t is a non-degenerate point. In particular, for any $t \in [0,T]$, $\widetilde{f}^{\epsilon}$ provides a P_N-regularization of f^0 away from degenerate points.

It is interesting to notice that according to (5.2), the evolved semi-classical Wigner function $\widetilde{f}^{\epsilon}$ is not only transported by the Hamiltonian flow but also dispersed in the $\mathbf{k}$-direction, through its convolution with the kernel $\tilde{G}^{\epsilon}$. Such a behavior is in complete agreement with the transport-dispersive character of the Wigner equation (cf. (1.12)).

Remark 5.1 (*Non essential potentials*). Let us see an interesting special case. We say that the potential is nonessential if $\mathbf{V}_k \equiv 0$, or equivalently, $V(\mathbf{x})$ is either zero, or linear in $\mathbf{x}$ or quadratic in $\mathbf{x}$. In this case the (full) Wigner equation is again a simple transport equation –it coincides with the Liouville equation satisfied by the limit Wigner f^0. Hence, the exact solution is given by $f^\epsilon(\mathbf{x},\mathbf{k},t) = f_0^\epsilon(\hat{\mathbf{q}},\hat{\mathbf{p}})$.

Let us see how $\widetilde{f}^\epsilon$ behaves. When $\mathbf{V}_k \equiv 0$, it follows from (3.16) that $\hat{\mathbf{U}}_k(t) \equiv 0$. Therefore, from (5.1) and (2.9) we have that $\tilde{G}^\epsilon(\mathbf{p},\hat{\mathbf{U}}_k(t)) = \tilde{G}^\epsilon(\mathbf{p},0) = \delta(\mathbf{p})$. Hence, (5.2) becomes

$$\widetilde{f}^\epsilon(\mathbf{x},\mathbf{k},t) = \delta(\mathbf{p}) *_{\mathbf{p}} \widetilde{W}_0^\epsilon(\hat{\mathbf{q}},\mathbf{p})\Big|_{\mathbf{p}=\hat{\mathbf{p}}} = \widetilde{W}_0^\epsilon(\hat{\mathbf{q}},\hat{\mathbf{p}}) = \widetilde{f}_0^\epsilon(\hat{\mathbf{q}},\hat{\mathbf{p}}).$$

That is, the evolution law for $\widetilde{f}^\epsilon$ coincides with that of f^ϵ. In addition, if $A_0 = constant$ and the initial phase $S_0(\mathbf{q})$ is a cubic polynomial in $\mathbf{q}$, then we easily see that $\widetilde{f}^\epsilon(\mathbf{x},\mathbf{k},t) \equiv f^\epsilon(\mathbf{x},\mathbf{k},t)$.

For nonessential potentials, non-degenerate points of the initial Lagrangian manifold are moved by the Hamiltonian flow to non-degenerate points. Indeed, at a non-degenerate point $(\mathbf{q},\nabla S_0(\mathbf{q}))$ on Λ_0 we have that $\sum_{i=1}^N(\boldsymbol{\sigma}^T\overline{\mathbf{C}}_i(0)\boldsymbol{\sigma})^2 = 0$ iff $\boldsymbol{\sigma} = 0$. At time t this point moves to the point $(\overline{\mathbf{x}}(t;\mathbf{q}),\ \nabla S(\overline{\mathbf{x}}((t;\mathbf{q}),t)$ on Λ_t and the third derivative matrices at this point are given by (all quantities evaluated along the ray $\overline{\mathbf{x}}(t;\mathbf{q})$)

$$\overline{\mathbf{C}}_k(t) = \overline{\Phi}(t)\overline{\mathbf{C}}_l(0)\overline{\Phi}^T(t)\overline{\phi}_{kl}(t).$$

The result then follows by the argument of Remark 1 of Appendix A.1 (with $\mathbf{T} = \overline{\Phi}^T(t)$).

Consequently, if the semi-classical Wigner function is originally a smooth function in a neighborhood of a point of the Lagrangian manifold, it will stay a smooth function at later times.

Remark 5.2 (*The $\epsilon = 0$ limit*). Let us take the limit $\epsilon \to 0$ in (5.2). Then $\widetilde{W}_0^\epsilon(\hat{\mathbf{q}},\hat{\mathbf{p}}) \to f_0^0(\hat{\mathbf{q}},\hat{\mathbf{p}})$, whereas by (A 6), $\tilde{G}^\epsilon(\mathbf{p},\hat{\mathbf{U}}_k(t)) \to \delta(\mathbf{p})$, as $\epsilon \to 0$. Consequently, as $\epsilon \to 0$, we have that

$$\widetilde{f}^\epsilon(\mathbf{x},\mathbf{k},t) = \tilde{G}^\epsilon(\mathbf{p},\hat{\mathbf{U}}_k(t)) *_{\mathbf{p}} \widetilde{W}_0^\epsilon(\hat{\mathbf{q}},\mathbf{p})\Big|_{\mathbf{p}=\hat{\mathbf{p}}} \to \delta(\mathbf{p}) *_{\mathbf{p}} f_0^0(\hat{\mathbf{q}},\mathbf{p})\Big|_{\mathbf{p}=\hat{\mathbf{p}}} = W_0^0(\hat{\mathbf{q}},\hat{\mathbf{p}}).$$

That is, in the limit $\epsilon = 0$, the dispersion mechanism disappears and we recover the Liouville equation satisfied by the limit Wigner f^0.

Remark 5.3 (*Approaching a caustic*). In order to arrive at the definition (5.2) we required that the evolved semi-classical Wigner be in agreement $\widetilde{W}^\epsilon$. In particular all the underlying analysis was restricted in the time interval $(0,T)$, $T < t_c$. Once however we define $\widetilde{f}^\epsilon$ by (5.2), then $\widetilde{f}^\epsilon$ is well defined even at a caustic point $(\mathbf{x}_c,t_c)$, if this point is a non-degenerate point. This is best seen in the case of nonessential potentials. At such a point, the integration $\int_{\mathbf{R}_{\mathbf{k}}^N}\widetilde{f}^\epsilon(\mathbf{x}_c,\mathbf{k},t_c)d\mathbf{k}$ is now meaningful.

6 A 2D example

To get some insight about the multiphase case, we will present an elementary but exact two-dimensional example of an elliptic umbilic caustic that evolves naturally from suitable initial data. We will consider equation (1.3) with $V(\mathbf{x}) \equiv 0$, and suitable WKB initial data. We will compute $|\psi^\epsilon(\mathbf{x},t)|$, at any point $(\mathbf{x},t)$ including the caustic points, by first finding $f^\epsilon(\mathbf{x},\mathbf{k},t)$ and then using the projection formula (1.2). The natural way to use (1.2) is in conjunction with suitable phase-space projection identities derived in [5]. We start with some general facts.

6.1 Lagrangian manifold–caustic

The initial manifold Λ_0 is the graph of a function from $\mathbf{R}^2_{\mathbf{q}}$ to $\mathbf{R}^2_{\mathbf{p}}$ and is given by $\Lambda_0 = \{(\mathbf{q},\mathbf{p}) : \mathbf{p} = \nabla S_0(\mathbf{q})\}$. At time t the phase flow moves Λ_0 to $\Lambda_t = \{(\mathbf{x},\mathbf{k}) : \hat{\mathbf{p}}(\mathbf{x},\mathbf{k},t) = \nabla_{\mathbf{q}} S_0(\hat{\mathbf{q}}(\mathbf{x},\mathbf{k},t))\}$. Whereas Λ_0 projects onto $\mathbf{R}^2_{\mathbf{q}}$ in a one-to-one way, this is not the case with Λ_t, $t>0$, for which there will be points of Λ_t, where the projection $(\mathbf{x},\mathbf{k}) \to \mathbf{x}$ is locally not invertible. The projection of these points onto $\mathbf{R}^2_{\mathbf{x}} \times \mathbf{R}_t$ defines the caustic in the physical variables (x_1,x_2,t).

We next derive the equations describing the caustic. The points of Λ_t which project onto the caustic should satisfy:

$$F_1(\mathbf{x},\mathbf{k},t) := \hat{p}_1(\mathbf{x},\mathbf{k},t) - \frac{\partial}{\partial q_1} S_0(\hat{\mathbf{q}}(\mathbf{x},\mathbf{k},t)) = 0,$$
$$F_2(\mathbf{x},\mathbf{k},t) := \hat{p}_2(\mathbf{x},\mathbf{k},t) - \frac{\partial}{\partial q_2} S_0(\hat{\mathbf{q}}(\mathbf{x},\mathbf{k},t)) = 0,$$

as well as

$$(\partial F_1/\partial k_1)(\partial F_2/\partial k_2) - (\partial F_1/\partial k_2)(\partial F_2/\partial k_1) = 0. \tag{6.1}$$

These are three equations involving five variables (k_1,k_2,x_1,x_2,t). Considering k_1, k_2 as parameters, they describe the caustic as a two dimensional surface in (x_1,x_2,t).

6.2 The zero potential

In the special case $V(x)=0$ the inverse bicharacteristics are given by:

$$\mathbf{q} = \hat{\mathbf{q}}(\mathbf{x},\mathbf{k},t) = \mathbf{x} - \mathbf{k}t, \qquad \mathbf{p} = \hat{\mathbf{p}}(\mathbf{x},\mathbf{k},t) = \mathbf{k}. \tag{6.2}$$

In this simple case it is convenient for subsequent calculations, to write the equations with respect to (q_1,q_2,x_1,x_2,t) instead of (k_1,k_2,x_1,x_2,t). To this end we note that by the chain rule we have

$$\frac{\partial F_i}{\partial k_j} = \frac{\partial \hat{p}_i}{\partial k_j} - \frac{\partial^2 S_0}{\partial q_i \partial q_1}\frac{\partial \hat{q}_1}{\partial k_j} - \frac{\partial^2 S_0}{\partial q_i \partial q_2}\frac{\partial \hat{q}_2}{\partial k_j}, \qquad i,j=1,2. \tag{6.3}$$

In addition,

$$k_i = \frac{(x_i - q_i)}{t}, \qquad \frac{\partial \hat{p}_i}{\partial k_j} = \delta_{ij}, \qquad \frac{\partial \hat{q}_i}{\partial k_j} = -t\delta_{ij}, \qquad i,j=1,2. \tag{6.4}$$

The equations describing the caustic are then given by:

$$\begin{aligned} \frac{x_1 - q_1}{t} - \frac{\partial}{\partial q_1} S_0(q_1, q_2) &= 0, \\ \frac{x_2 - q_2}{t} - \frac{\partial}{\partial q_2} S_0(q_1, q_2) &= 0, \\ \left(1 + t\frac{\partial^2 S_0}{\partial^2 q_1}\right)\left(1 + t\frac{\partial^2 S_0}{\partial^2 q_2}\right) - t^2\left(\frac{\partial^2 S_0}{\partial q_1\, \partial q_2}\right)^2 &= 0. \end{aligned} \tag{6.5}$$

Considering q_1, q_2 as parameters, the above three equations describe the caustic as a two dimensional surface in (x_1, x_2, t).

6.3 The elliptic umbilic caustic

We choose the WKB initial data $\psi_0(\mathbf{q}) = A_0(\mathbf{q})e^{\frac{i}{\epsilon}S_0(\mathbf{q})}$, with $A_0(\mathbf{q}) \equiv 1$ and

$$S_0(q_1, q_2) = \frac{1}{3}q_1^3 - q_1 q_2^2 - a(q_1^2 + q_2^2), \qquad a > 0. \tag{6.6}$$

Setting

$$u_1 := \frac{x_1}{t}, \qquad u_2 := \frac{x_2}{t}, \qquad v := \frac{1}{2t} - a, \tag{6.7}$$

we easily compute from (6.5) that the equations of the caustic are

$$\begin{aligned} u_1 &= q_1^2 - q_2^2 + 2q_1 v \\ u_2 &= -2q_1 q_2 + 2q_2 v \\ v^2 &= q_1^2 + q_2^2. \end{aligned} \tag{6.8}$$

For $v > 0$, we set $q_1 = v\cos\theta$, $q_2 = v\sin\theta$. We then compute

$$\begin{aligned} u_1 &= v^2(\cos 2\theta + 2\cos\theta) \\ u_2 &= -v^2(\sin 2\theta - 2\sin\theta). \end{aligned}$$

For fixed v (that is, fixed t) this describes a hypocycloid with three cusps. A similar analysis shows that the picture remains the same for $v < 0$. Returning to (x_1, x_2, t) space we obtain the elliptic umbilic caustic with a focus at the point $(x_1, x_2, t) = (0, 0, \frac{1}{2a})$.

The t=constant sections of the elliptic umbilic caustic are shown in Figure 2. The coordinates of the three cusps are easily found to be

$$\begin{aligned} (x_1, x_2)_A &= \left(3t\left(\frac{1}{2t} - a\right)^2, 0\right) \\ (x_1, x_2)_B &= \left(-\frac{3t}{2}\left(\frac{1}{2t} - a\right)^2, \frac{3\sqrt{3}t}{2}\left(\frac{1}{2t} - a\right)^2\right) \\ (x_1, x_2)_C &= \left(-\frac{3t}{2}\left(\frac{1}{2t} - a\right)^2, -\frac{3\sqrt{3}t}{2}\left(\frac{1}{2t} - a\right)^2\right). \end{aligned}$$

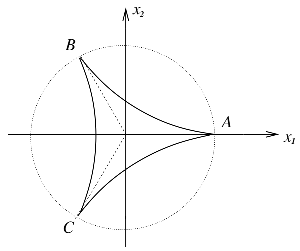

FIGURE 2. $t =$ constant sections of the elliptic umbilic caustic.

All three points move to infinity as t tends to either 0 or infinity. Also, at $t = 1/2a$, the hypocycloid reduces to a point (the origin) which is called the focus of the caustic.

6.4 Evolution of the scaled Wigner function (f^ϵ) and the wave field (ψ^ϵ)

The initial Wigner is given by

$$f_0^\epsilon(\mathbf{q}, \mathbf{p}) = \frac{1}{(\epsilon\pi)^2} \int_{\mathbf{R}^2} e^{\frac{i}{\epsilon} F(\mathbf{q},\mathbf{p};\boldsymbol{\sigma})} \, d\boldsymbol{\sigma}, \tag{6.9}$$

with

$$F(\mathbf{q}, \mathbf{p}; \boldsymbol{\sigma}) = -2(\mathbf{p} - \nabla S_0(\mathbf{q})) \cdot \boldsymbol{\sigma} + \frac{1}{3} \sum_{i,j,k=1}^{2} S_{0,q_i q_j q_k} \sigma_i \sigma_j \sigma_k.$$

The Wigner function at any time t is given by

$$f^\epsilon(\mathbf{x}, \mathbf{k}, t) = f_0^\epsilon(\mathbf{x} - \mathbf{k}t, \mathbf{k}) = f_0^\epsilon\left(\mathbf{q}, \frac{\mathbf{x} - \mathbf{q}}{t}\right), \qquad \left(\mathbf{k} = \frac{\mathbf{x} - \mathbf{q}}{t}\right).$$

We note that in this simple example, $f^\epsilon(\mathbf{x}, \mathbf{k}, t) = \widetilde{f}^\epsilon(\mathbf{x}, \mathbf{k}, t)$ (as defined in § 5). Now, a straightforward calculation shows that for S_0 as in (6.6)

$$F\left(\mathbf{q}, \frac{\mathbf{x} - \mathbf{q}}{t}; \boldsymbol{\sigma}\right) = -2\hat{\mathbf{z}} \cdot \boldsymbol{\sigma} + \frac{1}{3}(2\sigma_1^3 - 6\sigma_1\sigma_2^2),$$

where $\hat{\mathbf{z}} = (\hat{z}_1, \hat{z}_2)$

$$\hat{z}_1(\mathbf{u}, v; \mathbf{q}) := u_1 - 2vq_1 + q_2^2 - q_1^2,$$
$$\hat{z}_2(\mathbf{u}, v; \mathbf{q}) := u_2 - 2vq_2 + 2q_1q_2,$$

and $(\mathbf{u}, v)$ as defined by (6.7).

Throughout this section the phase integral P_2 is given by

$$P_2(\mathbf{z}) = \frac{1}{(2\pi)^2} \int_{\mathbf{R}^2} e^{i[\mathbf{z}\cdot\boldsymbol{\sigma} + \frac{1}{3}(2\sigma_1^3 - 6\sigma_1\sigma_2^2)]} \, d\boldsymbol{\sigma}.$$

We also define a new phase integral $\tilde{P}_2$ by

$$\tilde{P}_2(\mathbf{z}, w) := \frac{1}{(2\pi)^2} \int_{\mathbf{R}^2} e^{i[\mathbf{z}\cdot\boldsymbol{\sigma} + w|\boldsymbol{\sigma}|^2 + \frac{1}{3}(2\sigma_1^3 - 6\sigma_1\sigma_2^2)]}\, d\boldsymbol{\sigma}. \tag{6.10}$$

Notice that $\tilde{P}_2(\mathbf{z}, 0) = P_2(\mathbf{z})$. Moreover, for $\hat{\mathbf{z}}$ as above and $\gamma > 0$, the two phase integrals are related by the following "projection identity" (see Appendix C):

$$\int_{\mathbf{R}^2} P_2(-\gamma\hat{\mathbf{z}}(\mathbf{u}, v; \mathbf{q}))\, d\mathbf{q} = \frac{2^{1/3}4\pi^2}{\gamma} |\tilde{P}_2(-2^{-\frac{2}{3}}\gamma\mathbf{u},\ 2^{1/6}\gamma^{1/2}v)|^2. \tag{6.11}$$

Using our notation the Wigner function at any time t is given by

$$\begin{aligned} f^\epsilon(\mathbf{x}, \mathbf{k}, t) &= \frac{1}{(\epsilon\pi)^2} \int_{\mathbf{R}^2} e^{\frac{i}{\epsilon}(-2\hat{z}_1\sigma_1 - 2\hat{z}_2\sigma_2 + \frac{1}{3}(2\sigma_1^3 - 6\sigma_1\sigma_2^2)}\, d\boldsymbol{\sigma} \\ &= (2\epsilon^{-\frac{2}{3}})^2 P_2(-2\epsilon^{-\frac{2}{3}}\hat{\mathbf{z}}(\mathbf{u}, v; \mathbf{q})). \end{aligned}$$

On the other hand, the modulus of the amplitude is given by

$$\begin{aligned} |\psi^\epsilon(\mathbf{x}, t)|^2 &= \int_{\mathbf{R}^2} f^\epsilon(\mathbf{x}, \mathbf{k}, t) d\mathbf{k} \\ \left(\mathbf{k} = \frac{\mathbf{x} - \mathbf{q}}{t}\right) &:= \frac{1}{t^2} \int_{\mathbf{R}^2} f_0^\epsilon\left(\mathbf{q}, \frac{\mathbf{x} - \mathbf{q}}{t}\right) d\mathbf{q} \\ (\gamma = 2\epsilon^{-\frac{2}{3}}) &:= \frac{\gamma^2}{t^2} \int_{\mathbf{R}^2} P_2(-\gamma\hat{\mathbf{z}}(\mathbf{u}, v; \mathbf{q}))\, d\mathbf{q} \\ (\text{by } (6.11)) &:= \frac{2^{\frac{1}{3}}4\pi^2\gamma}{t^2} |\tilde{P}_2(-2^{-\frac{2}{3}}\gamma\mathbf{u},\ 2^{1/6}\gamma^{1/2}v)|^2. \end{aligned}$$

Recalling the definition of $\tilde{P}_2$ (cf. (6.10)), and changing variables by $\sigma_i = (2\epsilon)^{-1/3}\xi_i$, $(i = 1, 2)$, we have that

$$|\psi^\epsilon(\mathbf{x}, t)|^2 = \frac{1}{(2\pi\epsilon t)^2} \left| \int_{\mathbf{R}^2} e^{\frac{i}{\epsilon}(\frac{1}{3}\xi_1^3 - \xi_1\xi_2^2 + v(\xi_1^2 + \xi_2^2) - u_1\xi_1 - u_2\xi_2)} d\xi_1 d\xi_2 \right|^2, \tag{6.12}$$

where $\mathbf{u} = \frac{\mathbf{x}}{t}$ and $v = \frac{1}{2t} - a$. This formula is valid at any point including the points on the caustic. For instance, at the focus of the caustic $(x_1, x_2, t) = (0, 0, \frac{1}{2a})$ we just set in (6.12) $u_1 = u_2 = v = 0$. In particular, $|\psi^\epsilon(0, 0, \frac{1}{2a})| = O(\epsilon^{-1/3})$.

7 Conclusion

We have introduced the semiclassical Wigner function $\widetilde{f^\epsilon}$ as a formal approximation of the scaled Wigner trasform of the WKB solution to the problem (1.3)–(1.4). This approximation is valid near the manifold $\mathbf{k} = \nabla_\mathbf{x} S(\mathbf{x}, t)$. $\widetilde{f^\epsilon}$ is an object that is "richer" than the limit Wigner function f^0 which is a Dirac mass concentrated on $\mathbf{k} = \nabla_\mathbf{x} S(\mathbf{x}, t)$. In particular, $\widetilde{f^\epsilon}$ is an ϵ-dependent oscillatory integral that tends to f^0 as ϵ tends to zero. Moreover, it obeys an evolution law which is in agreement with the transport-dispersive character of the Wigner equation (1.12)–(1.13).

If $(\mathbf{x}_c, t_c)$ is a caustic point, the restriction of f^0 at this point, that is $f^0(\mathbf{x}_c, \mathbf{k}, t_c)$, is not a well defined distribution in $\mathbf{R}^N_{\mathbf{k}}$, and as a consequence the amplitude at $(\mathbf{x}_c, t_c)$ cannot be computed via the projection identity (1.17). On the contrary, $\widetilde{f}^\epsilon(\mathbf{x}_c, \mathbf{k}, t_c)$ is a well defined function in $\mathbf{R}^N_{\mathbf{k}}$, and therefore the integral of $\widetilde{f}^\epsilon$ with respect to $\mathbf{k}$ is meaningful and is expected to approximate the (ϵ-dependent) exact amplitude, at least on simple caustics. However, an approximation result in this direction is still missing.

More generally, the asymptotic nearness of $\widetilde{f}^\epsilon$ and f^ϵ is an open question. The difficulty in settling these questions stems from the fact that one deals with complicated multidimensional oscillatory integrals.

Acknowledgements

We thank the referee for several sharp comments and suggestions. We acknowledge partial support by the RTN European network HYKE, HPRN-CT-2002-00282. SF acknowledges additional support by the RTN European network Fronts–Singularities, HPRN-CT-2002-00274.

Appendix A The phase integral P_N

A.1 Regularity of P_N

Here we will make some remarks concerning the regularity of the phase integral P_N defined in (2.8). We recall that

$$P_N(\mathbf{z}, \mathbf{C}_k) = \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{i\phi(\mathbf{z},\boldsymbol{\sigma})}\, d\boldsymbol{\sigma}, \tag{A 1}$$

with

$$\phi(\mathbf{z}, \boldsymbol{\sigma}) := \mathbf{z}^T \boldsymbol{\sigma} + g(\boldsymbol{\sigma}), \qquad g(\boldsymbol{\sigma}) := \frac{1}{3} \boldsymbol{\sigma}^T \mathbf{C}_k \boldsymbol{\sigma} \sigma_k.$$

We note that P_N can be thought of as the Fourier transform of the $C^\infty(\mathbf{R}^N)$ function $e^{ig(\boldsymbol{\sigma})}$ and therefore is always well defined as a distribution. Also, since $g(\boldsymbol{\sigma})$ depends smoothly on the coefficients $\mathbf{C}_k$, the distribution P_N also depends smoothly on $\mathbf{C}_k$. We next find conditions under which P_N is a $C^\infty(\mathbf{R}^N)$ function.

We will follow closely the arguments of Hormander [12, §1.2], to which we refer for more details. Let us denote by $R = R(\mathbf{z})$, $C = C(\mathbf{z})$ two positive constants that may depend on $\mathbf{z}$. We then define the open set

$$Z_{\boldsymbol{\sigma}} := \{\mathbf{z} \in \mathbf{R}^N : \text{there exist } R,\ C \text{ s.t. for } |\boldsymbol{\sigma}| > R,\ \ |\nabla_{\boldsymbol{\sigma}} \phi(\mathbf{z}, \boldsymbol{\sigma})| \geqslant C|\boldsymbol{\sigma}|^2\}.$$

For $\mathbf{z} \in Z_{\boldsymbol{\sigma}}$ it follows easily that (i) the function $\boldsymbol{\sigma} \to \phi(\mathbf{z}, \boldsymbol{\sigma})$ has no critical points for $|\boldsymbol{\sigma}| > R$ and (ii) the quantity $\psi(\mathbf{z}, \boldsymbol{\sigma}) = |\nabla_{\boldsymbol{\sigma}} \phi(\mathbf{z}, \boldsymbol{\sigma})|^{-2}$ is a symbol with $\psi \in S^{-4}(\mathbf{R}^N)$, for large $|\boldsymbol{\sigma}|$; see Hormander [12, p. 83] for the definition of the symbol. Regarding then $\mathbf{z}$ as a parameter in the definition of P_N, we can use similar arguments as in Hormander [12], to show that $P_N(\mathbf{z}) \in C^\infty(Z_{\boldsymbol{\sigma}})$.

Let us see in particular the analogue of Lemma 1.2.1 in Hormander [12, p. 89]. The essential difference between this Lemma and our case is the fact that in [12] the phase is a symbol of order 1 whereas in our case the phase, due to the presence of the cubic terms, is a symbol of order 3.

Let U be a bounded open subset of $Z_{\boldsymbol{\sigma}}$. By the definition of $Z_{\boldsymbol{\sigma}}$ there exists an R_U such that $\nabla_{\boldsymbol{\sigma}}\phi(\mathbf{z},\boldsymbol{\sigma}) \neq 0$, for $|\boldsymbol{\sigma}| > R_U$ and $\mathbf{z} \in U$. Let $\chi(\boldsymbol{\sigma}) \in C_0^\infty(\mathbf{R}^N)$ be a smooth cutoff such that $\chi = 1$, for $|\boldsymbol{\sigma}| < R_U$. Regarding $\mathbf{z} \in U$ as a parameter, and starting from the identity

$$-\frac{i(1-\chi)}{|\nabla_{\boldsymbol{\sigma}}\phi|^2}\nabla_{\boldsymbol{\sigma}}\phi \cdot \nabla_{\boldsymbol{\sigma}}(e^{i\phi}) + \chi e^{i\phi} = e^{i\phi},$$

it is easy to see that the analogue of the operator L is given by

$$L = \mathbf{a} \cdot \nabla_{\boldsymbol{\sigma}} + c,$$

with

$$\mathbf{a} = (a_1, \ldots, a_N) = \frac{i(1-\chi)}{|\nabla_{\boldsymbol{\sigma}}\phi|^2}\nabla_{\boldsymbol{\sigma}}\phi, \qquad c = \chi + \mathrm{div}_{\boldsymbol{\sigma}}\mathbf{a}.$$

For $\mathbf{z} \in Z_{\boldsymbol{\sigma}}$, we have that $|\nabla_{\boldsymbol{\sigma}}\phi|^{-2} \in S^{-4}(\mathbf{R}^N)$, for large $|\boldsymbol{\sigma}|$ and it follows that $a_i \in S^{-2}(\mathbf{R}^N)$, $i = 1, \ldots N$, and $c \in S^{-3}(\mathbf{R}^N)$ for large $|\boldsymbol{\sigma}|$. The key observation here is that the coefficients a_i and c enjoy better decaying properties than in Lemma 1.2.1 in Hormander [12] – where $a_i \in S^0(\mathbf{R}^N)$ and $c \in S^{-1}(\mathbf{R}^N)$. This difference by 2, in the order of the symbols of the coefficients, is precisely what is needed to make up for the difference in the phases.

With this in mind, one can now argue exactly as in Hormander [12, pp. 89–90], to conclude that $P_N(\mathbf{z}) \in C^\infty(U)$. Since U is an arbitrary subset of $Z_{\boldsymbol{\sigma}}$, it follows that $P_N(\mathbf{z}) \in C^\infty(Z_{\boldsymbol{\sigma}})$.

We next show that if

$$|\nabla_{\boldsymbol{\sigma}} g(\boldsymbol{\sigma})|^2 = \sum_{i=1}^{N} (\boldsymbol{\sigma}^T \mathbf{C}_i \boldsymbol{\sigma})^2 \neq 0, \qquad \forall \boldsymbol{\sigma} \in \mathbf{R}^N \setminus \{0\}, \tag{A 2}$$

then $Z_{\boldsymbol{\sigma}} = \mathbf{R}^N$. Indeed, if $C_0 = \min_{|\boldsymbol{\sigma}|=1} |\nabla_{\boldsymbol{\sigma}} g(\boldsymbol{\sigma})| > 0$, by homogeneity we have that $|\nabla_{\boldsymbol{\sigma}} g(\boldsymbol{\sigma})| > C_0|\boldsymbol{\sigma}|^2$, for $\boldsymbol{\sigma} \neq 0$. On the other hand for any (fixed) $\mathbf{z} \in \mathbf{R}^N$ we have

$$|\nabla_{\boldsymbol{\sigma}}\phi(\mathbf{z},\boldsymbol{\sigma})| = \sum_{i=1}^{N} |z_i + \boldsymbol{\sigma}^T \mathbf{C}_i \boldsymbol{\sigma}| \geqslant \sum_{i=1}^{N} |\boldsymbol{\sigma}^T \mathbf{C}_i \boldsymbol{\sigma}| - \sum_{i=1}^{N} |z_i| \geqslant C_0|\boldsymbol{\sigma}|^2 - \sqrt{N}|\mathbf{z}| \geqslant \frac{C_0}{2}|\boldsymbol{\sigma}|^2,$$

for $|\boldsymbol{\sigma}|$ large enough. Hence $\mathbf{z} \in Z_{\boldsymbol{\sigma}}$, and consequently $Z_{\boldsymbol{\sigma}} = \mathbf{R}^N$.

We collect these observations in the following.

Proposition A.1 *The phase integral* P_N *in* (A 1) *is always well defined as a distribution and depends smoothly on the coefficients* $\mathbf{C}_k$. *If condition* (A 2) *is satisfied then* $P_N(\mathbf{z})$ *is a* $C^\infty(\mathbf{R}^N)$ *function.*

We next present some examples. For $N = 1$, $g(\sigma) = \frac{1}{3}c\sigma^3$ and condition (A 2) is satisfied iff $c \neq 0$, in which case $P_1(z)$ is the Airy function, whereas for $c = 0$, $P_1(z) = \delta(z)$.

For $N = 2$ any cubic form can be put, by a linear transformation, into one of the following five (nonequivalent) forms: (a) the hyperbolic case: $g(\boldsymbol{\sigma}) = \frac{1}{3}(\sigma_1^3 + \sigma_2^3)$, (b) the elliptic case: $g(\boldsymbol{\sigma}) = \frac{1}{3}\sigma_1^3 - \sigma_1\sigma_2^2$ (c) the parabolic case: $g(\boldsymbol{\sigma}) = \sigma_1\sigma_2^2$, (d) $g(\boldsymbol{\sigma}) = \frac{1}{3}\sigma_1^3$, (e) $g(\boldsymbol{\sigma}) = 0$; e.g. see Guillemin & Sternberg [10, Chapt. 7, Prop. 7.2]. Cases (a) and (b) are easily seen to satisfy condition (A 2) and therefore the corresponding P_2 integrals are $C^\infty(\mathbf{R}^2)$. The other three cases fail to satisfy condition (A 2). It is easy to check that for case (d) $P_2(z_1, z_2) = \mathrm{Ai}(z_1)\delta(z_2)$ whereas in case (e) $P_2(z_1, z_2) = \delta(z_1, z_2)$. In case (c) the corresponding integral can also be explicitly computed and is given by

$$P_2(z_1, z_2) = \begin{cases} \frac{\cos(z_2\sqrt{|z_1|})}{4\pi^2|z_1|}, & z_1 < 0, \\ 0, & z_1 > 0, \end{cases} \tag{A 3}$$

which is a C^∞ function in $\mathbf{R}^2 \setminus \{z_1 = 0\}$ (cf. Ben-Artzi *et al.* [1] and Fedoryuk [7]).

In all these examples $P_N(\mathbf{z})$, $N = 1, 2$ is in fact a signed measure with

$$\int_{\mathbf{R}^N} P_N(\mathbf{z})\, d\mathbf{z} = 1.$$

This is easily checked in all cases above, using the explicit forms of P_N.

We finally note that 2D oscillatory integrals with cubic phase have been studied in Fedoryuk [7] for analyzing Green functions for ultrahyperbolic equations, and recently in Ben-Artzi *et al.* [1] for deriving dispersion estimates for nonlinear Schrodinger-type systems.

A.2 Some properties of P_N

To avoid dealing with distributions we suppose that condition (A 2) is satisfied, so that P_N is a smooth function. The general case where P_N is a distribution is discussed in Remark 2 at the end of this section.

At first we note that

$$\int_{\mathbf{R}^N} P_N(\mathbf{z}, \mathbf{C}_k)\, d\mathbf{z} = 1. \tag{A 4}$$

At the formal level this follows by doing first the $d\mathbf{z}$-integration and then the $d\boldsymbol{\sigma}$-integration and using the fact

$$\frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{i\mathbf{z}\cdot\boldsymbol{\sigma}}\, d\mathbf{z} = \delta(\boldsymbol{\sigma}). \tag{A 5}$$

This can be made rigorous as follows. Let $\chi(\boldsymbol{\sigma}) \in C_0^\infty(B)$ be a smooth cutoff function supported in the unit ball B and such that $\int_{\mathbf{R}^N} \chi(\boldsymbol{\sigma})\, d\boldsymbol{\sigma} = 1$. We set $\chi_\epsilon(\boldsymbol{\sigma}) = \epsilon^{-N}\chi(\boldsymbol{\sigma}/\epsilon)$. and denote by $\check{\chi}_\epsilon(\mathbf{z})$ the inverse Fourier transform of $\chi_\epsilon(\boldsymbol{\sigma})$, that is

$$\chi_\epsilon(\boldsymbol{\sigma}) = \frac{1}{(2\pi)^N} \int_{\mathbf{R}^N} e^{i\mathbf{z}\cdot\boldsymbol{\sigma}} \check{\chi}_\epsilon(\mathbf{z})\, d\mathbf{z}.$$

We note that $\chi_\epsilon(\boldsymbol{\sigma}) \to \delta(\boldsymbol{\sigma})$ as $\epsilon \to 0$, whereas the function $\check{\chi}_\epsilon(\mathbf{z})$ decays rapidly at infinity and $\check{\chi}_\epsilon(\mathbf{z}) \to 1$, uniformly on compact sets, as $\epsilon \to 0$. We then have that

$$\int_{\mathbf{R}^N} P_N(\mathbf{z})\check{\chi}_\epsilon(\mathbf{z})\, d\mathbf{z} = \frac{1}{(2\pi)^N}\int_{(\mathbf{R}^N)^2} e^{i\mathbf{z}\cdot\boldsymbol{\sigma}}\check{\chi}_\epsilon(\mathbf{z})e^{ig(\boldsymbol{\sigma})}\, d\mathbf{z}\, d\boldsymbol{\sigma} = \int_{\mathbf{R}^N} \chi_\epsilon(\boldsymbol{\sigma})e^{ig(\boldsymbol{\sigma})}\, d\boldsymbol{\sigma},$$

and the result follows by sending ϵ to zero.

We next have that

$$\frac{1}{\eta^N}P_N\left(\frac{\mathbf{z}}{\eta}, \mathbf{C}_k\right) \to \delta(\mathbf{z}), \qquad \text{as } \eta \to 0. \tag{A 6}$$

To prove this we will show that if $\phi(\mathbf{z})$ is a $C_0^\infty(\mathbf{R}^N)$ function then

$$\frac{1}{(2\pi)^N}\frac{1}{\eta^N}\int_{(\mathbf{R}^N)^2} e^{i\frac{\mathbf{z}}{\eta}\cdot\boldsymbol{\sigma}}e^{ig(\boldsymbol{\sigma})}\phi(\mathbf{z})\, d\mathbf{z}\, d\boldsymbol{\sigma} \to \phi(0), \qquad \eta \to 0.$$

The left-hand side is easily seen to be equal to

$$\frac{1}{\eta^N}\int_{\mathbf{R}^N} \hat{\phi}(\boldsymbol{\sigma}/\eta)e^{ig(\boldsymbol{\sigma})}\, d\boldsymbol{\sigma} = \int_{\mathbf{R}^N} \hat{\phi}(\mathbf{x})e^{ig(\eta\mathbf{x})}\, d\mathbf{x} = \phi(0) + \int_{\mathbf{R}^N} \hat{\phi}(\mathbf{x})[e^{ig(\eta\mathbf{x})} - 1]\, d\mathbf{x},$$

where $\hat{\phi}$ denotes the Fourier transform of ϕ. Concerning the last integral we have that

$$\int_{\mathbf{R}^N} \hat{\phi}(\mathbf{x})[e^{ig(\eta\mathbf{x})} - 1]\, d\mathbf{x} \leqslant C\int_{|\mathbf{x}|<R} |e^{ig(\eta\mathbf{x})} - 1|\, d\mathbf{x} + 2\int_{|\mathbf{x}|>R} |\hat{\phi}(\mathbf{x})|\, d\mathbf{x} =: A + B.$$

The function $\hat{\phi}(\mathbf{x})$ decays at infinity faster than any power of $|x|$ and therefore we can make the term B smaller than, say, $\epsilon/2$ by taking R large enough. We then take η small enough to make the term A smaller than $\epsilon/2$ and the result follows.

Since P_N depends smoothly on the coefficients $\mathbf{C}_k$, we have that if $\mathbf{C}_k^\eta \to \mathbf{C}_k^0$ as $\eta \to 0$, then also

$$P_N(\mathbf{z}, \mathbf{C}_k^\eta) \to P_N(\mathbf{z}, \mathbf{C}_k^0), \qquad \text{as } \eta \to 0. \tag{A 7}$$

We next show a change of variables formula. If $\mathbf{T}$ is a nonsingular $N \times N$ matrix and

$$\tilde{\mathbf{z}} = \mathbf{T}^T\mathbf{z}, \qquad \tilde{\mathbf{C}}_k = \mathbf{T}^T\mathbf{C}_l\mathbf{T}T_{lk}, \qquad k, l = 1, \ldots, N, \tag{A 8}$$

we then have that

$$P_N(\mathbf{z}, \mathbf{C}_k) = |\det\mathbf{T}|\; P_N(\tilde{\mathbf{z}}, \tilde{\mathbf{C}}_k). \tag{A 9}$$

To prove this we notice that by the linear change of variables $\boldsymbol{\sigma} = \mathbf{T}\boldsymbol{\rho}$ in the integral (2.8), the value of the P_N stays the same. We then compute

$$\mathbf{z}^T\boldsymbol{\sigma} + \frac{1}{3}\boldsymbol{\sigma}^T\mathbf{C}_k\boldsymbol{\sigma}\sigma_k = (\mathbf{T}^T\mathbf{z})\boldsymbol{\rho} + \frac{1}{3}\boldsymbol{\rho}^T(\mathbf{T}^T\mathbf{C}_l\mathbf{T}T_{lk})\boldsymbol{\rho}\rho_k, = \tilde{\mathbf{z}}^T\boldsymbol{\rho} + \frac{1}{3}\boldsymbol{\rho}^T\tilde{\mathbf{C}}_k\boldsymbol{\rho}\rho_k.$$

Since $d\boldsymbol{\sigma} = |\det\mathbf{T}|\; d\boldsymbol{\rho}$, (A 9) follows from (A 1).

Let us note that the role of $\mathbf{z}$ in $\widetilde{W}^{\epsilon}$ is played by $(\mathbf{k} - \nabla_{\mathbf{x}} S(\mathbf{x}))$; cf. (2.10).

Then $\tilde{\mathbf{z}} = \mathbf{T}^T \mathbf{z}$, means $\tilde{\mathbf{k}} = \mathbf{T}^T \mathbf{k}$ and $\nabla_{\tilde{\mathbf{x}}} S(\tilde{\mathbf{x}}) = \mathbf{T}^T \nabla_{\mathbf{x}} S(\mathbf{x}))$. The last equality holds if $\tilde{\mathbf{x}} = T^{-1}\mathbf{x}$. Indeed, if $\mathbf{T} = \{T_{ij}\}$ then $S_{\tilde{x}_i}(\tilde{\mathbf{x}}) = S_{x_j}(\mathbf{x}) T_{ji}$. On the other hand by the chain rule $S_{\tilde{x}_i}(\tilde{\mathbf{x}}) = S_{x_j}(\mathbf{x}) \frac{\partial x_j}{\partial \tilde{x}_i}$. Hence $T_{ji} = \frac{\partial x_j}{\partial \tilde{x}_i}$, and $x_j = T_{ji} \tilde{x}_i$, or $\mathbf{x} = \mathbf{T}\tilde{\mathbf{x}}$.

The same transformation ($\mathbf{x} = \mathbf{T}\tilde{\mathbf{x}}$) yields the correct transformation of the third derivatives. Indeed, since $S_{\tilde{x}_i}(\tilde{\mathbf{x}}) = S_{x_l}(\mathbf{x}) T_{li}$, then

$$\begin{aligned}(\tilde{\mathbf{C}}_k)_{ij} = \tilde{c}_{ijk} = S_{\tilde{x}_i, \tilde{x}_j, \tilde{x}_k}(\tilde{\mathbf{x}}) &= S_{x_l, x_m, x_n}(\mathbf{x}) T_{li} T_{mj} T_{nk} = c_{lmn} T_{li} T_{mj} T_{nk} \\ &= (\mathbf{T}^T)_{il} (\mathbf{C}_n T)_{lj} T_{nk} = (\mathbf{T}^T \mathbf{C}_n \mathbf{T})_{ij} T_{nk},\end{aligned}$$

which is the same as (A 8). Thus, at the level of our original variables $(\mathbf{x}, \mathbf{k})$, the transformation (A 8) is equivalent to $(\tilde{\mathbf{x}}, \tilde{\mathbf{k}}) = (\mathbf{T}^{-1}\mathbf{x}, \mathbf{T}^T \mathbf{k})$.

We next derive a "convolution formula". This is essentially the well known fact that the Fourier transform of a convolution is the product of the Fourier transforms. Our starting point is the following formula valid for any $\boldsymbol{\alpha} \in \mathbf{R}^N$,

$$\int_{\mathbf{R}^N} e^{i\{(\mathbf{z}-\boldsymbol{\alpha})\cdot\boldsymbol{\sigma} + [g(\boldsymbol{\sigma}) + g'(\boldsymbol{\sigma})]\}} \, d\boldsymbol{\sigma} = \frac{1}{(2\pi)^N} \int_{(\mathbf{R}^N)^3} e^{i\{(\mathbf{z}-\boldsymbol{\alpha}-\mathbf{y})\cdot\mathbf{t} + g(\mathbf{t})\}} \, e^{i\{\mathbf{y}\cdot\boldsymbol{\rho} + g'(\boldsymbol{\rho})\}} \, d\mathbf{t} \, d\boldsymbol{\rho} \, d\mathbf{y}. \qquad \text{(A 10)}$$

To prove this, starting from the right hand side, we first integrate the $\mathbf{y}$-variable, taking into account that

$$\frac{1}{(2\pi)^N} \int_{\mathbf{R}} e^{i\mathbf{y}(\boldsymbol{\rho}-\mathbf{t})} d\mathbf{y} = \delta(\boldsymbol{\rho} - \mathbf{t}),$$

and (A 10) follows. To make rigorous this formal argument, one can use suitable approximation sequences, as in the proof of (A 4).

A direct consequence of (A 10) is

$$P_N(\mathbf{z} - \boldsymbol{\alpha}, \mathbf{C}_k + \mathbf{C}'_k) = P_N(\mathbf{z} - \boldsymbol{\alpha}, \mathbf{C}_k) *_{\mathbf{z}} P_N(\mathbf{z}, \mathbf{C}'_k). \qquad \text{(A 11)}$$

The following extension is easily seen to be true ($\eta \in \mathbf{R}$, $\eta \neq 0$):

$$P_N(\eta(\mathbf{z} - \boldsymbol{\alpha}), \mathbf{C}_k + \mathbf{C}'_k) = P_N((\eta(\mathbf{z} - \boldsymbol{\alpha}), \mathbf{C}_k) *_{\mathbf{z}} P_N(\eta \mathbf{z}, \mathbf{C}'_k) |\eta^N|. \qquad \text{(A 12)}$$

This relation is used in § 5.

Remark A.1 The nondegeneracy condition (A 2) is, of course, invariant under linear transformations. Indeed, if $\mathbf{T}$ is a nonsingular matrix and $\boldsymbol{\sigma} = \mathbf{T}\boldsymbol{\rho}$, it follows from (A 8) that

$$\sum_{i=1}^{N} (\boldsymbol{\rho}^T \tilde{\mathbf{C}}_i \boldsymbol{\rho})^2 = 0 \quad \Leftrightarrow \quad \sum_{i=1}^{N} (\boldsymbol{\sigma}^T \mathbf{C}_l \boldsymbol{\sigma} T_{li})^2 = 0.$$

The last equality is true iff $\boldsymbol{\sigma}^T \mathbf{C}_l \boldsymbol{\sigma} T_{li} = 0$ for all $i = 1, \ldots, N$, and this in its turn is equivalent to

$$(\boldsymbol{\sigma}^T \mathbf{C}_1 \boldsymbol{\sigma}, \ \boldsymbol{\sigma}^T \mathbf{C}_2 \boldsymbol{\sigma}, \ldots, \boldsymbol{\sigma}^T \mathbf{C}_N \boldsymbol{\sigma}) \cdot \mathbf{T} = 0 \quad \Leftrightarrow \quad \sum_{i=1}^{N} (\boldsymbol{\sigma}^T \mathbf{C}_i \boldsymbol{\sigma})^2 = 0.$$

Remark A.2 It is easy to see that, with the exception of (A 4), the above proofs with minor modifications, work even in the case where P_N is a distribution. Concerning (A 4), since in general we cannot integrate a distribution, one has first to prove that P_N is a (signed) measure. As explained in the first part of this appendix, this is true for $N = 1, 2$. We expect that it is true also for $N \geqslant 3$ but we do not have a proof.

Appendix B Some ODEs analysis

Here we will derive the solution of the matrix differential equation (3.14). We first consider the homogeneous problem, that is

$$\frac{d}{dt}\mathbf{C}_k(t) = -\mathbf{B}(t)\mathbf{C}_k(t) - \mathbf{C}_k(t)\mathbf{B}(t) - b_{kl}(t)\mathbf{C}_l(t), \qquad k, l = 1, 2, \ldots, N, \tag{B 1}$$

with $\mathbf{C}_k(0)$ given symmetric matrices. We recall that $\mathbf{B}(t)$ and $\mathbf{C}_k(t)$, $k = 1, 2, \ldots, N$ are symmetric $N \times N$ matrices with elements $b_{ij}(t)$ and $c_{ijk}(t)$, $i, j = 1, 2, \ldots, N$, respectively.

We have also denoted by $\Phi(t) = \{\phi_{ij}(t)\}$, $i, j = 1, 2, \ldots, N$ the fundamental solution of (3.4), that is

$$\frac{d}{dt}\Phi(t) = -\mathbf{B}(t)\Phi(t), \qquad \Phi(0) = \mathbf{I}_N,$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix.

We then look for the solution of (B 1) in the form

$$\mathbf{C}_k(t) = \Phi(t)\mathbf{M}_k(t)\Phi^T(t), \qquad k = 1, 2, \ldots, N, \tag{B 2}$$

for suitable matrices $\mathbf{M}_k(t)$. Notice in particular that $\mathbf{C}_k(0) = \mathbf{M}_k(0)$. If we plug (B 2) in (B 1) we see that the $\mathbf{M}_k(t)$'s satisfy the equations

$$\frac{d}{dt}\mathbf{M}_k(t) = -b_{kl}(t)\mathbf{M}_l(t), \qquad k, l = 1, 2, \ldots, N.$$

This can be written in matrix-block form as

$$\frac{d}{dt}\begin{pmatrix}\mathbf{M}_1(t)\\ \vdots \\ \mathbf{M}_2(t)\end{pmatrix} = -\begin{pmatrix} b_{11}(t)\mathbf{I}_N & \ldots & b_{1N}(t)\mathbf{I}_N \\ \vdots & \ddots & \vdots \\ b_{N1}(t)\mathbf{I}_N & \ldots & b_{NN}(t)\mathbf{I}_N\end{pmatrix}\begin{pmatrix}\mathbf{M}_1(t)\\ \vdots \\ \mathbf{M}_N(t)\end{pmatrix}. \tag{B 3}$$

The fundamental solution of (B 3) is the $N^N \times N^N$ matrix

$$\mathbf{X}(t) = \begin{pmatrix} \phi_{11}(t)\mathbf{I}_N & \ldots & \phi_{1N}(t)\mathbf{I}_N \\ \vdots & \ddots & \vdots \\ \phi_{N1}(t)\mathbf{I}_N & \ldots & \phi_{NN}(t)\mathbf{I}_N\end{pmatrix}, \qquad \mathbf{X}(0) = \mathbf{I}_{N^N}, \tag{B 4}$$

whence,

$$\mathbf{M}_k(t) = \phi_{kl}(t)\mathbf{M}_l(0) = \phi_{kl}(t)\mathbf{C}_l(0), \qquad k, l = 1, 2, \ldots, N.$$

It then follows from this and (B 2) that

$$\mathbf{C}_k(t) = \Phi(t)\,[\phi_{kl}(t)\mathbf{C}_l(0)]\,\Phi^T(t), \qquad k,l = 1,2,\dots,N. \tag{B 5}$$

We next consider the non-homogeneous case

$$\frac{d}{dt}\mathbf{C}_k(t) = -\mathbf{B}(t)\mathbf{C}_k(t) - \mathbf{C}_k(t)\mathbf{B}(t) - b_{kl}(t)\mathbf{C}_l(t) + \mathbf{F}_k(t), \qquad k,l = 1,2,\dots,N. \tag{B 6}$$

with $\mathbf{F}_k(t)$ $(k = 1,\dots,N)$ given symmetric matrices. Using the variation of constants method we look for a particular solution of (B 6) in the form

$$\mathbf{C}_k^{NH}(t) = \Phi(t)\,[\phi_{kl}(t)\mathbf{U}_l(t)]\,\Phi^T(t), \qquad k,l = 1,2,\dots,N. \tag{B 7}$$

for suitable matrices $\mathbf{U}_k(t)$ with $\mathbf{C}_k^{NH}(0) = \mathbf{U}_k(0) = 0$. If we plug this expression for $\mathbf{C}_k$ in (B 6) we end up with

$$\phi_{kl}\frac{d}{dt}\mathbf{U}_l = \Phi^{-1}\mathbf{F}_k\Phi^{-T}, \qquad \mathbf{U}_k(0) = 0, \qquad k,l = 1,2,\dots,N. \tag{B 8}$$

The solution then of (B 6) is the sum of (B 5) and (B 7), that is $(k,l = 1,2,\dots N)$,

$$\mathbf{C}_k(t) = \Phi(t)\,[\phi_{kl}(t)\mathbf{C}_l(0)]\,\Phi^T(t) + \Phi(t)\,[\phi_{kl}(t)\mathbf{U}_l(t)]\,\Phi^T(t). \tag{B 9}$$

Appendix C A phase-space projection identity

Here we will derive the projection identity (6.11).

$$\int_{\mathbf{R}^2} P_2(-\gamma\hat{\mathbf{z}}(\mathbf{u},v;\mathbf{q}))\,d\mathbf{q} = \frac{2^{1/3}4\pi^2}{\gamma}|\tilde{P}_2(-2^{-\frac{2}{3}}\gamma\mathbf{u},\ 2^{1/6}\gamma^{1/2}v)|^2 \tag{C 1}$$

This is a particular case of a general set of projection identities derived by Berry & Wright [5]. Let us recall a few things from Berry & Wright [5]. Let

$$\phi_E(S_1,S_2;C_1,C_2,C_3) := -C_1S_1 - C_2S_2 - C_3(S_1^2 + S_2^2) + S_1^3 - 3S_1S_2^2, \tag{C 2}$$

and define the phase integral ψ_E by

$$\psi_E(C_1,C_2,C_3) = \frac{1}{2\pi}\int e^{i\phi_E(S_1,S_2;C_1,C_2,C_3)}\,dS_1\,dS_2. \tag{C 3}$$

If

$$\tilde{C}_1 := 2^{2/3}(C_1 + 2C_3U_1 + 3(U_2^2 - U_1^2))$$
$$\tilde{C}_2 := 2^{2/3}(C_2 + 2C_3U_2 + 6U_1U_2),$$

then, the following projection identity holds (see Berry & Wright [5], relation (27)):

$$|\psi_E(C_1,C_2,C_3)|^2 = \frac{2^{1/3}}{\pi}\int_{\mathbf{R}^2}\psi_E(\tilde{C}_1,\tilde{C}_2,0)\,dU_1\,dU_2. \tag{C 4}$$

Let us rewrite this in terms of our notation. By setting first $S_i = \left(\frac{2}{3}\right)^{\frac{1}{3}} \sigma_i$, $(i = 1, 2)$ and then

$$U_i = \frac{\gamma^{\frac{1}{2}}}{2^{\frac{1}{2}} 3^{\frac{1}{3}}} q_i, \quad C_i = \frac{3^{\frac{1}{3}}}{2} \gamma u_i, \quad (i = 1, 2), \qquad C_3 = -3^{\frac{2}{3}} 2^{-\frac{1}{2}} \gamma^{\frac{1}{2}} v,$$

We have that

$$P_2(-\gamma \hat{\mathbf{z}}(\mathbf{u}, v; \mathbf{q})) = \left(\frac{3}{2}\right)^{\frac{2}{3}} \frac{1}{2\pi} \psi_E(\tilde{C}_1, \tilde{C}_2, 0),$$

and

$$\tilde{P}_2(-2^{-\frac{2}{3}} \gamma u_1,\ -2^{-\frac{2}{3}} \gamma u_2,\ 2^{1/6} \gamma^{1/2} v) = \left(\frac{3}{2}\right)^{\frac{2}{3}} \frac{1}{2\pi} \psi_E(C_1, C_2, C_3),$$

and (C 1) follows from (C 4).

We note that, as shown in Berry & Wright [5], similar identities hold for a large class of caustics, including the hyperbolic umbilic, the shallow tail and other cuspoid caustics.

References

[1] Ben-Artzi, M., Koch, H. and Saut, J.-C. (2003) Dispersion estimates for third order equations in two dimensions. *Comm. Partial Diff. Equ.* **28**(11–12), 1943–1974.

[2] Bensoussan, A., Lions, J. L. and Papanicolaou, G. (1978) *Asymptotic Analysis for Periodic Structures*. North-Holland.

[3] Berry, M. V. (1977) Semi-classical mechanics in phase space. A study of Wigner's function. *Proc. Phil. Trans. Roy. Soc.* **287**, 237–271.

[4] Berry, M. V. and Balazs, N. L. (1979) Evolution of semiclassical quantum states in phase space. *J. Phys. A*, **12**, 625–642.

[5] Berry, M. V. and Wright, F. J. (1980) Phase-space projection identities for diffraction catastrophes. *J. Phys. A* **13**, 149–160.

[6] Coddington, E. A. and Levinson, N. (1984) *Theory of Ordinary Differential Equations*. Krieger.

[7] Fedoryuk, M. V. (1977) *The saddle point method (Metod Perelava)*. Nauka.

[8] Filippas, S. and Makrakis, G. N. (2003) Semiclassical Wigner function and geometrical optics. *Multiscale Model. Simul.* **1(4)**, 674–710.

[9] Gerard, P., Markowich, P. A., Mauser, N. J. and Poupaud, F. (1977) Homogenization limits and Wigner transforms. *Comm. Pure Appl. Math.* **50**, 323–380.

[10] Guillemin, V. and Sternberg, S. (1977) *Geometric Asymptotics*. American Mathematical Society.

[11] Haake, F. (2001) *Quantum Signatures of Chaos*. Springer.

[12] Hormander, L. (1971) Fourier integral operators I. *Acta Mathematica*, **127**, 79–183.

[13] Jin, S. and Li, X. (2003) Multi-phase computations of the semiclassical limit of the Schrödinger equation and related problems: Whitham vs. Wigner. *Phys. D*, **182**(1–2), 46-85.

[14] Kravtsov, Yu. A. (1968) Two new asymptotic methods in the theory of wave propagation in inhomogeneous media (review) *J. Sov. Phys. Acoust*. **14**(1), 1–17.

[15] Lax, P. D. (2002) *Functional Analysis*. Wiley Interscience.

[16] Lions, P. L. and Paul, T. (1993) Sur les measures de Wigner. *Rev. Math. Iberoamericana*, **9**, 563–618.

[17] Maslov, V. P. and Fedoryuk, V. M. (1981) *Semiclassical Approximations in Quantum Mechanics.* Reidel.

[18] Ozorio de Almeida, A. M. (1983) The Wigner function for two dimensional tori: Uniform approximation and projections. *Ann. Phys*, **145**, 100–115.

[19] Papanicolaou, G. and Ryzhik, L. (1999) Waves and transport. In: L. Cafarrelli (ed.), *Hyperbolic Equations and Frequency Interactions*, pp. 305–382. AMS.

[20] Rios, P. P. de M. and Ozorio de Almeida, A. M. (2002) On the propagation of semiclassical Wigner functions *J. Phys. A: Math. Gen.* **35**, 2609–2617.

[21] Sparber, C., Markowich, P. A. and Mauser, N. J. (2003) Wigner functions versus WKB-methods in multivalued geometrical optics *Asymptot. Anal.* **33**, 153–187.

[22] Tappert, F. (1977) The parabolic approximation method. In: J. B. Keller and J. S. Papadakis (eds.), *Wave Propagation and Underwater Acoustics,* pp. 224–287. Lecture Notes in Physics 70. Springer-Verlag.

[23] Tatarskii, V. (1971) *The Effects of the Turbulent Atmosphere in Wave Propagation.* Israel Program for Scientific Translation.

[24] Tatarskii, V. (1984) The Wigner representation of quantum mechanics. *Sov. Phys. Usp.*, **26**, 311–327.