Metaplectic geometrical optics for modeling caustics in uniform and non-uniform media

N A Lopez\(^1\)\(\dagger\) and I Y Dodin\(^1,2\)\(\dagger\)

\(^1\)Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, United States of America
\(^2\)Princeton Plasma Physics Laboratory, Princeton, NJ 08543, United States of America

E-mail: nlopez@pppl.gov

Received 7 September 2020, revised 19 November 2020
Accepted for publication 8 December 2020
Published 12 January 2021

Abstract

As an approximate theory that is highly regarded for its computational efficiency, geometrical optics (GO) is widely used for modeling waves in various areas of physics. However, GO fails at caustics, which significantly limits its applicability. A new framework, called metaplectic geometrical optics (MGO), has recently been developed that allows caustics of certain types to be modeled accurately within the GO framework. Here, we extend MGO to the most general case. To illustrate our new theory, we also apply it to several sample problems, including calculations of two-dimensional wavefields near fold and cusp caustics. In contrast with traditional GO solutions, the corresponding MGO solutions are finite everywhere and approximate the true wavefield well near these caustics.

Supplementary material for this article is available online

Keywords: metaplectic transform, caustics, catastrophe optics, geometrical optics, symplectic transformation, phase space

(Some figures may appear in colour only in the online journal)

1. Introduction

The propagation of waves in homogeneous and weakly inhomogeneous media is often described within the approximate theory known as geometrical optics (GO), or ray optics [1, 2]. However, GO fails at so-called caustics, where it predicts spurious singularities of the wavefield. Loosely speaking, caustics are surfaces across which the number of rays arriving at a given point changes abruptly [3]. The general properties of such surfaces have long been known from catastrophe theory, which provides a classification wherein only a finite number of caustic types is possible for a given number of spatial dimensions [4, 5]. This result underlies modern research into caustics [6–9], as general properties of a given caustic type can be inferred by studying a particular case. Still, practical calculations continue to rely on directly solving wave equations [10–12], which is computationally expensive.

It would be advantageous to find a more efficient way to calculate these caustic structures. In particular, the question of whether caustics can be modeled by somehow extending GO has been attracting attention for a long time.

One known solution to this problem is by Maslov \([13]\), who proposed to rotate the ray phase space occasionally by π/2 using the Fourier transform (FT) in one or more spatial variables. Such rotations can remove caustics and locally reinstate GO, but they are inconvenient for simulations because the rotation points have to be introduced \textit{ad hoc}, requiring the simulations to be supervised. In references \([14, 15]\), we proposed a modification of Maslov’s approach by replacing the FT with the metaplectic transform (MT), which might be familiar from other contexts such as aberration theory \([16]\) and paraxial optics \([17–19]\). Using the MT, one can transform the wavefield continually along the rays, both eliminating the singularities at caustics and allowing for fully automated and
fast simulations. This new framework, which we call metaplectic geometrical optics (MGO), has already been successfully benchmarked on one-dimensional (1D) problems [15]. However, MGO as originally formulated in [14, 15] can still yield singularities in certain situations; thus, further progress is needed.

Here, we report a more general version of MGO, where the last spurious singularities are removed at the expense of additional calculations. We then demonstrate the reformulated MGO analytically on three examples: a homogeneous plane wave, a fold caustic, and a cusp caustic. In Arnold’s classification [20], they correspond to the A1, A2, and A3 caustic types, respectively. The first case has no caustic per se; it is considered only to present a basic tutorial on the MGO machinery. The two other cases show how MGO leads to solutions that, unlike traditional GO solutions, are finite everywhere and approximate the true wavefield near the caustics well.

This paper is organized as follows. In section 2, we review the basic equations of GO and MGO as described in [15]. In section 3, we report a more general version of MGO. In section 4, we discuss three examples of the reformulated MGO, and we show analytically that MGO adequately approximates the wavefield in the entire space, even at caustics. In section 5, we summarize our main conclusions. Auxiliary calculations are presented in the appendices and in the supplementary material (stacks.iop.org/JOPT/23/025601/mmedia).

2. Overview of MGO

We start by briefly reviewing the MGO method as it was presented in [15]. Let us consider a non-driven linear wave in a general linear medium. We do not assume any particular wave equation in this paper, as our theory is sufficiently general to handle any linear wave equation. However, for simplicity, we restrict our consideration to scalar equations. (Generalization to vector waves on non-Euclidean spaces is possible using the machinery presented in [21].)

Specifically, we assume that the wave is described by a scalar field \( \psi \) and governed by an integral equation

\[
\int dq' D(q, q') \psi(q') = 0. \tag{1}
\]

Here, \( q \) is the coordinate on an N-D Euclidean (or pseudo-Euclidean) space, called \( q \)-space, and \( D(q, q') \) is the dispersion kernel. In particular, differential wave equations (partial or ordinary) have dispersion kernels that consist of delta functions and their derivatives. For example, the Helmholtz equation has \( D(q, q') = \nabla^2 \delta(q' - q) + n^2(q') \delta(q' - q) \), where \( \nabla^2 \) is the gradient with respect to \( q' \) and \( n(q) \) is a spatially varying index of refraction (see section 4 for more examples).

More generally, the kernel \( D \) can be a smooth function, as is the case for waves in warm plasma, for example. (A review of the general theory of linear dispersion can be found in [22].)

It is not necessary for our purposes to specify this function; let us simply state that \( D \) encodes all information about the linear medium where the wave propagates, whatever that medium may be.

Consider also \( p \)-space, which is the Fourier-dual of \( q \)-space. Then, \( q \)-space and \( p \)-space collectively define a 2N-D phase space with coordinates \( (q, p) \). The dispersion kernel \( D(q, q') \) can be associated with a function on phase space via the transformation

\[
D(q, p) = \int d\mathbf{q} e^{i\bar{p} \cdot \mathbf{q}} D \left( \frac{q}{2}, \mathbf{q} + \frac{\mathbf{p}}{2} \right), \tag{2}
\]

where the integral is taken over \( q \)-space and \( \bar{\mathbf{p}} \) denotes definitions. (Note that vectors are interpreted as row vectors unless explicitly transposed via \( \tau \), so \( \mathbf{p} \tau = \mathbf{p} \cdot \).) The function \( D(q, p) \) is known as the Weyl symbol of \( D(q, q') \). We shall now analyze (1) in the short-wavelength limit using both traditional GO and MGO. Specifically, we assume in the following that the medium parameters, the wave envelope, and the wavelength itself vary slowly over a characteristic wavelength [1, 2, 21].

2.1. Traditional GO

In traditional GO, \( \psi(q) \) is assumed to take the eikonal form given by

\[
\psi(q) = \phi(q) e^{i\theta(q)}, \tag{3}
\]

where \( \phi(q) \) is a slowly varying envelope and \( \theta(q) \) is a rapidly varying phase. Then, to the lowest order, (1)–(3) yield the local dispersion relation [2, 21]

\[
D(q, k(q)) = 0, \tag{4}
\]

along with the envelope transport equation

\[
v(q) \tau \partial_q \log \phi(q) = -\frac{1}{2} \partial_q \cdot v(q). \tag{5}
\]

Here, the local wavevector \( k(q) \) and the local group velocity \( v(q) \) are defined as

\[
k(q) \equiv \partial_q \theta(q), \quad v(q) \equiv \partial_p D(q, p) \big|_{p=k(q)}. \tag{6}
\]

Hence, \( k(q) \) is irrotational, meaning that

\[
\partial_{q_m} k_{\ell} = \partial_{q_\ell} k_m, \quad \ell, m = 1, \ldots, N. \tag{7}
\]

Equations (4) and (5) are commonly solved along the characteristic rays that satisfy

\[
\partial_\tau q(\tau) = \partial_k D[q(\tau), k(\tau)], \tag{8a}
\]

\[
\partial_\tau k(\tau) = -\partial_q D[q(\tau), k(\tau)]. \tag{8b}
\]

For integrable systems, these rays trace out an N-D surface in phase space called the dispersion manifold. We then define \( \tau \equiv (\tau_1, \ldots, \tau_N) \) as coordinates on the dispersion manifold, with \( \tau_1 \) serving as the longitudinal coordinate along a ray. Initial conditions provide \( q(0, \tau_{\perp}) \) and \( k(0, \tau_{\perp}) \), where
\[ \tau_\perp = (\tau_2, \ldots, \tau_N). \] For example, \( \tau_1 \) can be the time variable or one of the spatial coordinates; then \( \tau_\perp \) are the remaining spatial coordinates.

The wavevector \( k(q) \) is determined from the rays as
\[
k(q) = k[\tau(q)],
\]
where \( \tau(q) \) is the function inverse to \( q(\tau) \). Although (8a) and (8b) show that rays cannot cross in phase space, their projections onto \( q \)-space can. As a result, \( \tau(q) \) is generally multi-valued, meaning \( k(q) \) is as well. The envelope is constructed along the rays as
\[
\phi_i(q(\tau)) = \phi_i(q(0, \tau_\perp)) \left[ \frac{j(0, \tau_\perp)}{j(\tau)} \right],
\]
where \( \phi_i[q(0, \tau_\perp)] \) is set by initial conditions, and
\[
j(\tau) \doteq \det \partial_\tau q(\tau).
\]

Then, the total wavefield is constructed as
\[
\psi(q) = \sum_{t \in \tau(q)} \phi_i[q(t)] \exp \left[ i \int dq^\top k(q) \right],
\]
where the summation is taken over all branches of \( \tau(q) \).

Clearly, \( \psi(q) \) diverges where \( j(t) = 0 \). Such points are caustics, and by (11), they occur where the dispersion manifold has a singular projection onto \( q \)-space. To extend GO modeling to caustics and the neighboring regions, Lopez et al [15] proposed MGO, which we now describe.

2.2. MGO

MGO also uses the rays provided by (8a) and (8b) to solve (4). However, in MGO, the phase space is not fixed but instead is continually rotated, \( (q, p) \rightarrow (Q, P) \), such that the local projection of the dispersion manifold onto \( Q \)-space is always well-behaved. Accordingly, the envelope equation (5) is replaced with a similar envelope equation in the rotated frame that has no caustics by construction. This is done as follows.

Let us assume that the dispersion manifold \( (q, p) = (q(\tau), k(\tau)) \) has been obtained by integrating (8a) and (8b) and consider the tangent plane at \( \tau = t \). We can rotate our original phase space to align \( q \)-space with the tangent plane at \( t \) (\( Q \)-space) using the following linear transformation:
\[
Q_i = A_i q_i + B_i p_i, \quad P_i = C_i q_i + D_i p_i.
\]

where \( P_i \) is the Fourier dual to \( Q_i \) and the matrices \( A_i, B_i, C_i, \) and \( D_i \) are all \( N \times N \). In (13), we require that
\[
S_i \doteq \begin{pmatrix} A_i & B_i \\ C_i & D_i \end{pmatrix}
\]
be symplectic, that is,
\[
S_i J S_i^\top = J_i = \begin{pmatrix} 0_N & I_N \\ -I_N & 0_N \end{pmatrix},
\]
where \( 0_N \) and \( I_N \) are respectively the \( N \times N \) null and identity matrices. (A practical algorithm for computing \( S_i \) from the ray trajectories using Gram–Schmidt orthogonalization is provided in [15].)

The transformation of the wavefield corresponding to the symplectic transformation (14) of the ray phase space is the MT [23, 14], sometimes called the linear canonical transform. The MT is a linear integral transformation from \( \psi \) to a new function \( \Psi \) given explicitly as
\[
\Psi(Q) = \int dq M(Q, q) \psi(q),
\]
where the MT kernel is given as
\[
M(Q, q) \doteq \frac{\sigma \exp [iG(Q, q)]}{(2\pi)^{N/2} \sqrt{\det B}},
\]
and \( G(Q, q) \) is the following quadratic phase function:
\[
G(Q, q) = \frac{1}{2} Q^\top DB^{-1}Q - Q^\top B^{-1}q + \frac{1}{2} q^\top B^{-1}Aq.
\]

(Here, \( -^\top \) denotes the matrix inverse transpose.) Each \( S \) actually has two corresponding MTs that differ by an overall sign, which is designated by \( \sigma \doteq \pm 1 \). Note that (17) requires \( \det B \neq 0 \); this requirement will be lifted in section 3.1. Also note that the MT reduces to the familiar FT when \( S = J \).

Let \( \Psi_i(Q) \) be the MT of \( \psi(q) \) corresponding to \( S_i \) of (14) (for a chosen MT sign convention). In MGO, we assume that \( \Psi_i(Q_i) \) [not \( \psi(q) \)] has the eikonal form
\[
\Psi_i(Q_i) \doteq \Phi_i(Q_i) e^{\Theta_i(Q_i)},
\]
where \( \Phi_i \) is a slowly varying complex envelope and \( \Theta_i \) is a phase that varies rapidly with the new coordinate \( Q_i \); hence, \( K_i(Q_i) \doteq \partial_{Q_i} \Theta_i(Q_i) \) is understood as the local wavevector in \( Q_i \)-space. The eikonal approximation (19) is facilitated by the fact that \( \partial_{Q_i} K_i = -\partial_{Q_i} K_i = 0 \) at the tangent point by definition, where
\[
\tilde{D}[Q_i, K_i(Q_i)] = D[D_i^\top Q_i - B_i^\top P_i, -C_i^\top Q_i + A_i^\top P_i]
\]
is the Weyl symbol of the dispersion operator in the new coordinates. (We also assume that \( \partial_{Q_i} K_i \) is slowly varying in the neighborhood of the tangent point.)

As shown in [15], taking the MT of (1) and performing the standard GO procedure in \( Q_i \)-space using (19) yields equations similar to (4) and (5) in the new variables. Specifically, the local dispersion relation has the form
\[
\tilde{D}[Q_i, K_i(Q_i)] = 0,
\]
and the envelope transport equation becomes
\[
V_i(Q_i)^\top \partial_{Q_i} \log \Phi_i(Q_i) = -\frac{1}{2} \partial_{Q_i} V_i(Q_i).
\]

Here, we have introduced
\[
V_i(Q_i) \doteq \partial_{Q_i} \tilde{D}[Q_i, P_i] \bigg|_{P_i = K_i(Q_i)},
\]
which is the corresponding group velocity.

If (4) is satisfied along the original rays, then (21) is satisfied along the rotated rays given by

\[ Q_{\alpha}(\tau) = A_{\alpha}(q(\tau)) + B_{\alpha}(k(\tau)) \tag{24a} \]

\[ K_{\alpha}(\tau) = C_{\alpha}(q(\tau)) + D_{\alpha}(k(\tau)) \tag{24b} \]

Hence, the dispersion manifold simply rotates with the ambient phase space³.

Before solving (22), it is convenient to renormalize \( \Psi_t(Q_0) \) by its value at \( Q_0(t) \); that is, let

\[ \Psi_t(Q_0(t)) = \alpha_t \Phi_t(Q_0(t)) e^{i\Theta_t(Q_0(t))}, \quad \alpha_t = \Psi_t(Q_0(t)) \tag{25} \]

and require

\[ \Phi_t[Q_0(t)] = 1, \quad \Theta_t[Q_0(t)] = 0. \tag{26} \]

Then, analogous to (10), (22) is solved to yield

\[ \Phi_t[Q_0(t)] = \Phi_t[Q_0(t, \tau_{\perp})] \left[ \frac{J(t_1, \tau_{\perp})}{J(t)} \right] \tag{27} \]

where \( \Phi_t[Q_0(t, \tau_{\perp})] \) is set by initial conditions subject to (26), and

\[ J(t) = \det \partial_\tau Q_0(t). \tag{28} \]

The phase can also be immediately determined as

\[ \Theta_t(Q_0) = \int_{Q_0(t)}^{Q_0} dQ^T K_t(Q), \tag{29} \]

where the line integral is taken over any path with the specified endpoints, and \( K_t(Q_0) \) is constructed as

\[ K_t(Q_0) = K_t[\tau(Q_0)], \tag{30} \]

where \( \tau(Q_0) \) is the function inverse to \( Q_0(\tau) \). Note that \( K_t(Q_0) \) will generally be multivalued, so it must be restricted to the branch satisfying \( K_t[Q_0(t)] = K_t(t) \).

By continuity, \( \alpha_t \) evolves along the rays as

\[ \alpha_t = \alpha_t(0, t_{\perp}) \exp \left[ \int_0^{t_{\perp}} dh \eta(h, t_{\perp}) \right], \tag{31} \]

where \( \alpha_t(0, t_{\perp}) \) is determined by initial conditions and

\[ \eta \equiv \frac{i}{2} K_t(t) W_t K_t(t) - \frac{i}{2} Q^T_t(t) U_t Q_t(t) - \frac{1}{2} \text{Tr} (V_t) + \left[ \partial_\tau Q_t(t) - V_t^T Q_t(t) - W_t^T K_t(t) \right]^T \]

\[ \times \left\{ \partial_\tau \Phi_t[Q_t(t)] + i K_t(t) \right\}. \tag{32} \]

Here, the \( N \times N \) matrices \( U_t, V_t, \) and \( W_t \) are defined via

\[ (\partial_\tau S_t) S_t^{-1} = \begin{pmatrix} V_t^T & W_t \\ -U_t & -V_t \end{pmatrix}, \tag{33} \]

where the directional derivative, defined as \( h \partial_\tau = h^T \partial_\tau \), should be interpreted as a total derivative acting on arguments and subscripts containing \( t \). (An alternate algorithm for evolving \( \alpha_t \) based on successive applications of a near-identity MT [14] is provided in [15].)

Applying an inverse MT then maps \( \Psi_t(Q_0) \) to a function on \( q \)-space, denoted \( \psi_t(q) \), as

\[ \psi_t(q) = \int dQ_0 M_t^{-1}(q, Q_0) \Psi_t(Q_0), \tag{34} \]

where the inverse MT kernel is given as

\[ M_t^{-1}(q, Q_0) = \frac{\sigma_t \exp \left[ -i G_t(q, Q_0) \right]}{(-2\pi i)^{N/2} \sqrt{\det B_t}}. \tag{35} \]

The overall sign factor \( \sigma_t \) must change whenever \( \det B_t \) crosses the branch cut of the square root to maintain continuity. This change in \( \sigma_t \) is related to the discrete phase jumps a wavefield experiences upon traversing a caustic.

At this point, semiclassical methods like GO and Maslov’s method traditionally evaluate (34) using the stationary-phase approximation (SPA) about the ray contribution \( Q_0 = Q_0(t) \) [25]. However, the SPA fails when saddlepoints are close together [26], as occurs near caustics. To remedy this, note that under fairly general conditions, integrals like (34) can be evaluated on the union of steepest-descent contours through some subset of saddlepoints in complex \( Q_0 \)-space [27]. By integrating (34) only along the steepest-descent contour through \( Q_0 = Q_0(t) \) rather than the entire set, we can isolate the desired ray contribution in a manner that is asymptotically equivalent to the SPA but is also well-behaved at caustics. (In this regard, we can also define the “saddlepoint contribution” to an integral as the result of integrating along the corresponding steepest-descent contour.) Hence, the variable shift \( \epsilon \equiv Q_0 - Q_0(t) \) yields

\[ \psi_t(q) = \frac{\sigma_t \alpha_t \exp \left[ -\frac{i}{2} \beta_t(q) \right]}{(-2\pi i)^{N/2} \sqrt{\det B_t}} Y_t(q), \tag{36} \]

where we have defined

\[ \beta_t(q) \equiv 2 G_t[q, Q_t(t)], \tag{37a} \]

\[ Y_t(q) = \int_{c_0} \exp \left[ -i \gamma_t(q, \epsilon, q) \right] \tag{37b} \]

\[ \gamma_t(q, \epsilon, q) \equiv \frac{1}{2} \epsilon^T D_t B_t^{-1} \epsilon + \epsilon^T B_t^{-1} [D_t^2 Q_t(t) - q]. \tag{37c} \]

and \( c_0 \) is the steepest-descent contour through \( \epsilon = 0 \). Finally, \( \psi(q) \) is reconstructed by summing over all branches of the dispersion manifold:

\[ \psi(q) = \sum_{t \in \tau(q)} \psi_t[q(t)]. \tag{38} \]
Equation (38) can accurately model the wavefield incident on an isolated fold caustic or bounded between a pair of fold caustics in 1D. However, (38) cannot model a wavefield whose dispersion manifold has \( \det \mathcal{B}_t = 0 \) over a finite domain (examples of which are discussed in section 4). We call such ray patterns 'quasi-uniform'. To enable MGO to model quasi-uniform ray patterns, the restriction that \( \det \mathcal{B}_t \neq 0 \) must be lifted. The corresponding theory is discussed in the next section.

3. MGO for quasi-uniform ray patterns

3.1. Singular MTs

The MT kernel \( M(Q, q; S) \) that corresponds to a symplectic matrix \( S \) with \( \det \mathcal{B} = 0 \) can be considered as a limiting case [23, 14]

\[
M(Q, q; S) = \lim_{\varepsilon \to 0} M(Q, q; S_\varepsilon),
\]

where \( S_\varepsilon \) is a symplectic matrix that has \( S \) as a limit at \( \varepsilon \to 0 \). (Here and further, we omit the subscript \( t \) for ease of notation.) For example, we can adopt

\[
S_\varepsilon = \begin{pmatrix}
a & B + \varepsilon A \\
b & C + \varepsilon D
\end{pmatrix},
\]

whose symplecticity (to all orders in \( \varepsilon \)) can be readily verified by definition (15).

To show that \( \det (B + \varepsilon A) \neq 0 \) and subsequently compute the limit in (39), it is useful to perform a singular value decomposition (SVD) of \( B \). Let \( \rho \) and \( \varsigma \equiv N - \rho \) be the rank and core-rank of \( B \), respectively. Then, the SVD of \( B \) takes the form

\[
B = L \tilde{B} R^\top,
\]

where \( \tilde{B} \) is a diagonal matrix given by

\[
\tilde{B} = \begin{pmatrix}
\Lambda_{pp} & 0_{\rho\varsigma} \\
0_{\varsigma\rho} & 0_{\varsigma\varsigma}
\end{pmatrix},
\]

(the submatrices with subscript \( mn \) are size \( m \times n \)) and \( \Lambda_{pp} \) is a diagonal matrix that has all nonzero singular values of \( B \) on its diagonal:

\[
\Lambda_{pp} = \begin{pmatrix}
\lambda_1 & & \\
& \ddots & \\
& & \lambda_\rho
\end{pmatrix}.
\]

Note that \( \det \Lambda_{pp} \neq 0 \) by definition. The matrices \( L \) and \( R \) are both orthogonal and can be written as

\[
L = \begin{pmatrix}
\uparrow & \uparrow & \cdots & \uparrow \\
\hat{e}_1 & \hat{e}_2 & \cdots & \hat{e}_N
\end{pmatrix}, \quad R = \begin{pmatrix}
\uparrow & \uparrow & \cdots & \uparrow \\
f_1 & f_2 & \cdots & f_N
\end{pmatrix}.
\]

The columns of these matrices are, respectively, the left singular vectors \( \{ \hat{e}_j \} \) and right singular vectors \( \{ f_j \} \) of \( B \), which are mutually orthonormal:

\[
\hat{e}_j^\top \hat{e}_k = \delta_{jk}, \quad f_j^\top f_k = \delta_{jk}.
\]

Let us similarly define

\[
\tilde{A} \doteq L^\top A R, \quad \tilde{C} \doteq L^\top C R, \quad \tilde{D} \doteq L^\top D R.
\]

As shown in appendix A, these matrices have the form

\[
\tilde{A} = \begin{pmatrix}
a_{pp} & a_{pc} \\
a_{cp} & a_{cc}
\end{pmatrix},
\]

\[
\tilde{C} = \begin{pmatrix}
c_{pp} & c_{pc} \\
c_{cp} & c_{cc}
\end{pmatrix},
\]

\[
\tilde{D} = \begin{pmatrix}
d_{pp} & 0_{pc} \\
d_{cp} & a_{cc}
\end{pmatrix}.
\]

Note that \( a_{cc} \) is invertible. Hence, we compute

\[
\det (B + \varepsilon A) = \det \begin{pmatrix}
\Lambda_{pp} + \varepsilon a_{pp} & \varepsilon a_{pc} \\
0_{\rho\varsigma} & \varepsilon a_{cc}
\end{pmatrix}
\]

\[
= \det (\Lambda_{pp} + \varepsilon a_{pp}) \det (\varepsilon a_{cc})
\]

\[
\approx \varepsilon^\rho \det \Lambda_{pp} \det a_{cc},
\]

where we have used \( \det L = \det R = 1 \). Since \( \det a_{cc} \neq 0 \) and \( \det \Lambda_{pp} \neq 0 \) by definition, \( \det (B + \varepsilon A) \) is nonzero for finite \( \varepsilon \). (We adopt the convention that \( 0 \times 0 \) matrices have unit determinant.)

By (17) and (48), we obtain to leading order in \( \varepsilon \)

\[
M(Q, q; S_\varepsilon) \approx \frac{\sigma e^{-\varepsilon^2/2} \exp \left[ ig(q_p, Q) \right]}{(2\pi)^{N/2} \sqrt{\det \Lambda_{pp} \det a_{cc}}} \times \exp \left( \frac{i}{2\varepsilon} \left| q_c - a_{cc}^{-1} Q_c \right|^2 \right),
\]

where we have defined

\[
g(q_p, Q) = \frac{1}{2} \left( q_p^\top \Lambda_{pp}^{-1} a_{pp} q_p - q_p^\top M_1 L^\top Q \right)
\]

\[
+ \frac{1}{2} Q^\top L M_2 L^\top Q,
\]

along with the matrices

\[
M_1 \doteq \begin{pmatrix}
\Lambda_{pp}^{-1} - \Lambda_{pp}^{-1} a_{pc} a_{cc}^{-1}
\end{pmatrix},
\]

\[
M_2 \doteq \begin{pmatrix}
d_{pp} \Lambda_{pp}^{-1} a_{pc} a_{cc}^{-1} \\
d_{cp} \Lambda_{pp}^{-1} a_{cc}^{-1}
\end{pmatrix},
\]

and the vector decompositions

\[
R^\top q = \begin{pmatrix}
q_p \\
q_c
\end{pmatrix}, \quad L^\top Q = \begin{pmatrix}
Q_p \\
Q_c
\end{pmatrix}.
\]
Note that $M_3$ and $\Lambda_{\rho \rho}^{-1} a_{\rho \rho}$ are symmetric ((A4), (A7), and (A16)). Also note that $M_1$ is size $\rho \times N$, $M_2$ is size $N \times N$, and any vector $\nu_\rho$ is size $m \times 1$ (appendix B provides details for the derivation of (49)). Then, using (39) along with \[
abla \lim_{\varepsilon \to 0} \varepsilon^{-\gamma/2} \exp \left( i \varepsilon \frac{\delta}{2} (q_\rho - a_{\rho \rho}^{-1} Q_\rho) \right) = (2\pi i)^{\gamma/2} \delta (q_\rho - a_{\rho \rho}^{-1} Q_\rho), \]
we obtain the limit of the MT kernel at $\det B \to 0$:
\[
M(Q, q) = \frac{\sigma \exp \left[ i g(Q, q) \right]}{(2\pi i)^{\gamma/2} \sqrt{\det \Lambda_{\rho \rho} \det a_{\rho \rho}^{-1}}},
\]
where, for brevity, we no longer mention the dependence of $M$ on the symplectic matrix explicitly.

Following straightforward delta-function manipulations, we obtain the inverse MT kernel when $\det B = 0$:
\[
M^{-1}(q, Q) = \frac{\sigma \exp \left[ -i g(Q, q) \right]}{(2\pi i)^{\gamma/2} \sqrt{\det \Lambda_{\rho \rho} \det a_{\rho \rho}^{-1}}},
\]
where we have defined
\[
\tilde{g}(Q, q) = \frac{1}{2} q^\top d_{\rho \rho} \Lambda_{\rho \rho}^{-1} Q_\rho - Q_\rho^\top M_3 R^\top q + \frac{1}{2} q^\top R M_2 R^\top q,
\]
along with the matrices
\[
M_1 \doteq \left( \Lambda_{\rho \rho}^{-1} - \Lambda_{\rho \rho}^{-1} d_{\rho \rho}^\top a_{\rho \rho} \right),
\]
\[
M_3 \doteq \left( \Lambda_{\rho \rho}^{-1} a_{\rho \rho}^\top \Lambda_{\rho \rho}^{-1} a_{\rho \rho}^\top \Lambda_{\rho \rho}^{-1} a_{\rho \rho} \right) - \left( \Lambda_{\rho \rho}^{-1} d_{\rho \rho}^\top a_{\rho \rho} \right) \Lambda_{\rho \rho}^{-1} a_{\rho \rho}.
\]

Note that $M_4$ and $d_{\rho \rho} \Lambda_{\rho \rho}^{-1} a_{\rho \rho}$ are symmetric ((A4), (A7), and (A16)). Also, $M_3$ has size $\rho \times N$ while $M_4$ has size $N \times N$. Lastly, we choose the following branch-cut convention: $\arg(i) = \pi/2$ and $\arg(\det \Lambda_{\rho \rho} \det a_{\rho \rho}^{-1}) \in (-\pi, \pi]$ in (54); $\arg(-i) = -\pi/2$ and $\arg(\det \Lambda_{\rho \rho} \det a_{\rho \rho}^{-1}) \in (-\pi, \pi]$ in (55).

3.2. Singular MGO

Let us now incorporate the general representation for the inverse MT given by (55) into the MGO formalism. We emphasize that (55) is valid for all values of $\det B$. Using (34), (52), and (55) yields\[
\psi_t(q) = N_t(q) \int dQ_\rho dQ_\varsigma \delta (Q_\rho - a_{\rho \rho}, Q_\varsigma) \times \exp \left( -i \frac{1}{2} Q_\rho^\top d_{\rho \rho} \Lambda_{\rho \rho}^{-1} Q_\rho + i Q_\rho^\top M_3 R^\top q \right) \times \psi_t \left[ \begin{bmatrix} Q_\rho \\ Q_\varsigma \end{bmatrix} \right],
\]
where we have substituted $Q$ with $Q_\rho$ and $Q_\varsigma$ as
\[
Q = L \begin{bmatrix} Q_\rho \\ Q_\varsigma \end{bmatrix}, \quad dQ = d(LQ) = dQ_\rho dQ_\varsigma
\]
and defined the prefactor
\[
N_t(q) \doteq \frac{\sigma t \exp \left( -i \frac{1}{2} q^\top R M_3 R^\top q \right)}{(2\pi i)^{\gamma/2} \sqrt{\det \Lambda_{\rho \rho} \det a_{\rho \rho}^{-1}}}
\]
(As a reminder, all matrices depend on $t$.) The integration over $Q_\varsigma$ is immediately performed to yield
\[
\psi_t(q) = N_t(q) \int dQ_\rho \psi_t \left[ \begin{bmatrix} Q_\rho \\ a_{\rho \rho}, Q_\varsigma \end{bmatrix} \right] \times \exp \left( -i \frac{1}{2} Q_\rho^\top d_{\rho \rho} \Lambda_{\rho \rho}^{-1} Q_\rho + i Q_\rho^\top M_3 R^\top q \right). \]

The phase of the integrand is stationary where
\[
\partial_{Q_\rho} \Theta \left[ \begin{bmatrix} Q_\rho \\ a_{\rho \rho}, Q_\varsigma \end{bmatrix} \right] + M_3 R^\top q - d_{\rho \rho} \Lambda_{\rho \rho}^{-1} Q_\rho = 0.
\]
When $q$ is evaluated at the ray location $q(t)$ in (38), then, using (8a) and (8b), we can simplify\[
M_3 R^\top q(t) = M_3 R^\top d^\top \Theta Q_\rho(t) - M_3 R^\top d^\top \Theta K_i(t)
\]
where we have defined the vector projections
\[
L^\top K_i(t) \doteq \begin{bmatrix} K_i^\rho(t) \\ K_i^\varsigma(t) \end{bmatrix},
\]
\[
L^\top Q_i(t) \doteq \begin{bmatrix} Q_i^\rho(t) \\ Q_i^\varsigma(t) \end{bmatrix} = \begin{bmatrix} Q_\rho(t) \\ a_{\rho \rho}, Q_\varsigma(t) \end{bmatrix}.
\]
Note that (64b) follows from (8a) to (8b). Since
\[
\partial_{Q_i} \Theta_i \left[ \begin{bmatrix} Q_\rho \\ a_{\rho \rho}, Q_\varsigma \end{bmatrix} \right] = K_i^\rho \left[ \begin{bmatrix} Q_\rho \\ a_{\rho \rho}, Q_\varsigma \end{bmatrix} \right]
\]
the saddlepoint criterion (62) therefore becomes
\[
\left\{ K_i^\rho \left[ \begin{bmatrix} Q_\rho \\ a_{\rho \rho}, Q_\varsigma(t) \end{bmatrix} \right] - K_i^\rho(t) \right\} + d_{\rho \rho} \Lambda_{\rho \rho}^{-1} [Q_i^\rho(t) - Q_\rho] = 0,
\]
where we have used the fact that $d_{\rho \rho} \Lambda_{\rho \rho}^{-1}$ is symmetric.
As can be verified, the desired point $t$ on the dispersion manifold is a root to (66), since both terms in brackets vanish simultaneously when $Q_\rho = Q_i^\rho(t)$. Let us therefore define the new integration variable
\[
\epsilon_\rho \doteq Q_\rho - Q_i^\rho(t), \quad d\epsilon_\rho = dQ_\rho.
\]
This yields a modified version of (36):

\[ \psi_\tau(q) = \frac{\sigma_\tau \alpha_\tau \exp \left[ -\frac{i}{2} \beta_\tau(q) \right]}{(-2\pi i)^{N/2} \sqrt{\det \Lambda_{\rho \rho} \det \Lambda_{\xi \xi}}} \} \Gamma_\tau(q), \]  

(68)

where we have defined

\[ \beta_\tau(q) = 2\alpha_\tau (Q_\tau(q), q), \]  

(69a)

\[ \Gamma_\tau(q) = \int_{C_0} d\epsilon_\tau \psi_\tau \left[ L \left( Q_\tau(q) + \epsilon_\tau \right) \right] \times \exp \left[ -i \epsilon_\tau(q) \right], \]  

(69b)

\[ \gamma_\tau(q) = \frac{1}{2} \epsilon_\tau(q) [d_{\rho \rho} \Lambda_{\rho \rho}^{-1} \epsilon_\tau(q) + \epsilon_\tau(q)]. \]  

(69c)

Note that \( \Gamma_\tau(q) \) is integrated along the steepest-descent contour \( C_0 \) passing through \( \epsilon_\tau = 0 \). Equation (38) with \( \psi_\tau \) computed via (68) constitutes the generalization of MGO to all values of \( \det B \). Thus, MGO can now be applied to any ray pattern in arbitrary media.

3.3. MGO with Gaussian coherent states

Instead of performing an SVD of \( B \), we can develop an expression equivalent to (68) using

\[ f(Q, Z_0) = \exp \left[ \left( -\frac{|Q - Q_0|^2}{2} - iK_0 (Q - Q_0) \right) \right]. \]  

(70)

These functions, which satisfy the completeness relation

\[ \delta(Q - Q') = \int \frac{dQ_0 dK_0}{(2\pi)^N} f(Q, Z_0) f(Q', Z_0)^*, \]  

(71)

can be understood as the spatial representations of the Gaussian coherent states centered around \( Z_0 \). These states are commonly used in quantum optics and are discussed in detail in appendix C. As shown in appendix C, the property (71) ultimately leads to an alternate representation of the MT:

\[ M_\tau^{-1}(q, Q) = \int dK_0 \frac{\sigma \exp \left[ G_t(q, \xi) - |K_0|^2 \right]}{(\sqrt{2\pi})^N \sqrt{\det (2D_1 - iB_1)}}, \]  

(72)

where we have defined \( \xi \equiv Q + 2iK_0 \) and

\[ G_t(q, \xi) = -\frac{1}{2} q^T (2D_1 - iB_1)^{-1} (A_1 + 2iC_1) q + \left( q - \frac{1}{2} D_1^T \xi \right)^T (2D_1 - iB_1)^{-1} \xi. \]  

(73)

Note that the complex matrix \( 2D - iB \) is always invertible [29].

With (72) as the MT kernel, the phase of (34) is stationary where \( Q \) and \( K_0 \) simultaneously satisfy

\[ q - D_1^T Q + B_1^T K_0 = 0, \]  

(74a)

\[ 2D_1^T [K_t(Q) - K_0] + q - D_1^T Q + B_1^T K_t(Q) = 0. \]  

(74b)

When \( q \) is evaluated at \( q_t(t) \), a simultaneous solution to (74a) and (74b) is \( Q = Q_t(t) \) and \( K_0 = K_t(t) \). Therefore, upon defining the new integration variables

\[ \epsilon_r \equiv Q - Q_t(t), \]  

\[ \epsilon_i \equiv 2K_0 - 2K_t(t), \]  

(75)

we obtain the following alternate representation of \( \psi_t(q) \):

\[ \psi_t(q) = \frac{\sigma_\tau \alpha_\tau \exp \left[ \tilde{G}_t[q, \xi_t(t)] - |K_t(t)|^2 \right]}{(2\sqrt{2\pi})^N \sqrt{\det (2D_1 - iB_1)}} \times \int_{C_0} d\epsilon_r d\epsilon_i \psi_t[\epsilon_r + Q_t(t)] \exp \left[ -\tilde{\gamma}_t(\epsilon, q, t) \right], \]  

(76)

where we have defined

\[ \epsilon \equiv \epsilon_r + i\epsilon_i. \]  

(77a)

\[ \xi_t(t) \equiv Q_t(t) + 2iK_t(t). \]  

(77b)

\[ \tilde{\gamma}_t(\epsilon, q, t) \equiv \frac{1}{2} \epsilon^T D_1 (2D_1 - iB_1)^{-1} \epsilon - \frac{1}{4} \epsilon^T [D_1 - |K_t(t)|^2] \]  

(77c)

Equation (76) is equivalent to (68) but might be advantageous since it can be applied 'as is' without performing an SVD of \( B \). That said, (76) involves a 2N-D integral, which is harder to evaluate numerically. In this sense, the representation (68) may be more practical, especially at large \( N \).

4. Examples

Here, we consider several examples of MGO, which have been shortened for clarity. The complete calculations can be found in the supplementary material.

4.1. Plane wave in uniform medium: no caustic

As a first example, let us consider a plane wave propagating in a uniform medium. For simplicity, we consider 1-D propagation governed by the one-way wave equation

\[ i\partial_q \psi(q) + \psi(q) = 0. \]  

(78)

There is no caustic in this case, and (78) is easy to integrate even without using MGO. However, this example is instructive to illustrate the reformulated MGO machinery when \( \det B_t = 0 \) with relatively little algebra.
Let us start by writing (78) in the integral form (1). The corresponding kernel \( D(q, q') \) can be written as
\[
D(q, q') = i\partial_q\delta(q - q') - \delta(q - q'),
\] (79)
so the Weyl symbol (2) is as follows:
\[
\mathcal{D}(q, k) = k - 1.
\] (80)
The corresponding ray equations are
\[
\partial_q q(\tau) = 1, \quad \partial_q k(\tau) = 0,
\] (81)
with solutions given by
\[
q(\tau) = \tau, \quad k(\tau) = 1,
\] (82)
where the integration constants have been chosen to satisfy \( \mathcal{D}[q(0), k(0)] = 0 \). Since \( \tau(q) = q \) is single-valued, \( \psi(q) \) will be absent of caustics.

From (81), it is clear that the tangent plane of the dispersion manifold is \( \mathbf{q} \)-space itself, that is,
\[
S_t = I_2
\] (83)
for all \( t \in \tau(q) \). Hence, \( \Psi_t(Q_t) \) is trivially obtained as
\[
\Psi_t(Q_t) = \alpha_t \exp(iQ_t - it),
\] (84)
where we have used \( Q_t(t) = t \). Since \( \partial_t S_t = 0_2 \), \( \partial_t \Psi_t(Q_t) = 0 \), and \( \partial_t Q_t(t) = 1 \), we also compute
\[
\eta_t = i, \quad \alpha_t = \alpha_0 \exp(it),
\] (85)
where \( \alpha_0 \) is a constant. Then, since \( B_t = 0 \) implies that \( \rho = 0 \), \( \varsigma = 1 \), and \( R = L = 1 \), then \( \beta_t^R = \gamma_t^R = 0 \) and the integration over \( d\rho \) is empty. Hence, (68) becomes
\[
\psi_t[q(t)] = \sigma_t \alpha_0 \exp[iQ_t(t)],
\] (86)
Since the branch cut of the MT is never crossed, we can take \( \sigma_t = 1 \). Then, the summation over branches is trivially performed to yield
\[
\psi(q) = \sum_{t \in \tau(q)} \psi_t[q(t)] = \alpha_0 \exp(iq).
\] (87)
Equation (87) is an exact solution of (78), which is anticipated because (78) is a first-order equation and thus coincides with its GO approximation.

### 4.2. Plane wave in linearly stratified medium: fold caustic

As a second example, let us consider oblique propagation in a linearly stratified medium. Suppose that the wave is described by the Helmholtz-type equation
\[
\partial_q^2 \psi(q) + (k_0^2 - q_1)\psi(q) = 0,
\] (88)
where \( k_0 \) is a constant and \( \partial_q^2 = \partial_q^2 + \partial_{q_1}^2 \). Our coordinate system is such that \( q_1 \) is aligned with the medium stratification and \( q_2 \) is transverse to \( q_1 \). Let us also consider the initial condition
\[
\psi(0, q_2) = c \exp(ik_0 q_2),
\] (89)
where \( c \) is an arbitrary constant.

Equation (88) can be equivalently written as an integral equation (1) with integration kernel
\[
D(q, q') = -\partial_q^2 \delta(q - q') + (q_1 - k_0^2)\delta(q - q'),
\] (90)
The corresponding Weyl symbol (2) is
\[
\mathcal{D}(q, k) = k_1^2 + k_2^2 + q_1 - k_0^2,
\] (91)
and the corresponding ray equations are
\[
\partial_{\tau_1} q_1(\tau) = 2k_1(\tau), \quad \partial_{\tau_1} k_1(\tau) = -1, \quad \partial_{\tau_1} k_2(\tau) = 0.
\] (92)
Let us define \( \tau_1 \) such that \( q_1(0, \tau_2) = 0 \). Then, the initial condition (89) implies that \( k_2(0, \tau_2) = k_0 \), and the local dispersion relation \( \mathcal{D}[q_1(0, \tau_2), k_1(0, \tau_2)] = 0 \) requires \( k_1(0, \tau_2) = 0 \). This leaves \( q_2(0, \tau_2) \) undetermined. Since \( \tau_2 \) must parameterize the initial conditions for the rays, let us choose \( q_2(0, \tau_2) = \tau_2 \).

Hence, the ray solutions are
\[
q_1(\tau) = -\tau_1^2, \quad q_2(\tau) = \tau_2 + 2k_0 \tau_1, \quad k_1(\tau) = -\tau_1, \quad k_2(\tau) = k_0.
\] (93)
The inverse function \( \tau(q) \) is calculated as
\[
\tau_1(q) = \pm \sqrt{-q_1}, \quad \tau_2(q) = q_2 + 2k_0 \sqrt{-q_1}.
\] (94)
Clearly, \( \tau(q) \) is double-valued, so there are two branches that must ultimately be summed over.

A basis for \( Q_1 \)-space is provided by the vector pair
\[
\partial_{\tau_1} z(t) = \begin{pmatrix} -2\tau_1 & 2k_0 & -1 & 0 \end{pmatrix}^T,
\] (95a)
\[
\partial_{\tau_2} z(t) = \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix}^T,
\] (95b)
where \( z(t) = (q(t), k(t))^T \). Then, symplectic Gram–Schmidt orthogonalization [15] yields the submatrices
\[
A_0 = D_0 = \frac{1}{\partial_{\tau_1}} \begin{pmatrix} -2\tau_1 & 0 & 0 \end{pmatrix},
\] (96a)
\[ B_t = -C_t = \frac{1}{\vartheta_t} \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \]  

where we have defined

\[ \vartheta_t \doteq \sqrt{1 + 4r_t^2}. \]  

The rotated rays are calculated via (8a) and (8b) as

\[ Q_t(\tau) = \frac{1}{\vartheta_t} \left( \frac{2n_t^2 + \tau_1}{\vartheta_t \tau_2 + 2 \vartheta_t k_t \tau_1} \right), \tag{98a} \]

\[ K_t(\tau) = \frac{1}{\vartheta_t} \left( \frac{2n_t \tau_1 - \tau_1^2}{\vartheta_t k_0} \right). \tag{98b} \]

We can therefore compute the inverse function \( \tau(Q_t) \), which is double-valued. Upon restricting \( K_t(Q_t) \) to the correct branch, we obtain

\[ K_{1.1}(Q_t) = -\frac{Q_t}{2r_t} - \vartheta_t \frac{1 - \sqrt{1 + 8r_t \vartheta_t Q_t}}{8r_t^2}, \tag{99a} \]

\[ K_{1.2}(Q_t) = k_0. \tag{99b} \]

Then, the line integral of (29) is computed to yield

\[ \Theta_t[\varepsilon + Q_t(\tau)] = k_0 c_2 + \frac{8r_t^4 + \vartheta_t^4}{8r_t^2} \vartheta_t \varepsilon_1 + \frac{1}{4r_t^2} \varepsilon_1 \]

\[ + \left( \vartheta_t^4 + 8r_t \vartheta_t \varepsilon_1 \right)^{3/2} - \vartheta_t^4 \]  

\[ + \frac{\varepsilon_1}{96r_t^3}. \tag{100} \]

Using (27) and (28), we next compute

\[ \Phi_t[\varepsilon + Q_t(\tau)] = \frac{\vartheta_t}{(\vartheta_t^4 + 8r_t \vartheta_t \varepsilon_1)^{1/4}}, \tag{101} \]

where we have chosen the initial conditions to satisfy

\[ \Phi_t[Q_t(t_1, \tau_2)] = 1. \tag{102} \]

We then compute via (31) and (32)

\[ \alpha_t = \frac{\alpha_t(0,0)}{\sqrt{\vartheta_t}} \exp \left( 2ik_0^2 t_1 + \frac{2r_t^3}{3} - \frac{i r_t^5}{\vartheta_t^4} \right), \tag{103} \]

where \( \alpha_t(0,0) \) is an arbitrary initial condition.

We now perform the inverse MT. Note that \( B_t \) is already in the desired SVD form, with \( L = R = l_2 \) and \( \rho = 1 \). Since \( \varepsilon_\rho = \varepsilon_1 \) and \( Q_t(\tau) = Q_{t,1}(t) \), we compute

\[ \gamma_t^0[\varepsilon_{\rho}, Q_{t}(\tau)] = \frac{r_t^2}{\vartheta_t} \varepsilon_1 + t_1 \varepsilon_1. \tag{104} \]

Hence, we obtain

\[ \gamma_t^0[\varepsilon_{\rho}, Q_{t}(\tau)] = \int_{c_2} d\varepsilon_1 \frac{\vartheta_t}{(\vartheta_t^4 + 8r_t \vartheta_t \varepsilon_1)^{1/4}} \]

\[ \times \exp \left[ \frac{(\vartheta_t^4 + 8r_t \vartheta_t \varepsilon_1)^{3/2} - \vartheta_t^4}{96r_t^3} - \frac{i \vartheta_t^4}{4r_t^3} \varepsilon_1 - \frac{i \vartheta_t^4}{8r_t^3} \varepsilon_1 \right]. \tag{105} \]

This is the same integral that was studied in [15], where the following approximation was derived:

\[ \gamma_t^0[\varepsilon_{\rho}, Q_{t}(\tau)] \approx \pi \vartheta_t \exp \left( -\frac{i}{3} r_t^6 \vartheta_t^4 \right) \]

\[ \times \left[ Ai \left( -\frac{r_t^2 \vartheta_t^4}{4} \right) - i \frac{t_1}{|t_1|} Bi \left( -\frac{r_t^2 \vartheta_t^4}{4} \right) \right]. \tag{106} \]

where \( Ai(x) \) and \( Bi(x) \) are the Airy functions of the first and second kind, respectively [30]. Thus, (68) yields

\[ \psi_t[\varepsilon_{\rho}, Q_{t}(\tau)] = i \sigma_t(0,0) \frac{\vartheta_t \sqrt{\pi}}{\sqrt{-2r_t^2}} \exp \left[ \frac{2ik_0^2 t_1 + \frac{2r_t^3}{3} (1 - \vartheta_t^4) \right] \]

\[ \times \left[ Ai \left( -\frac{r_t^2 \vartheta_t^4}{4} \right) - i \frac{t_1}{|t_1|} Bi \left( -\frac{r_t^2 \vartheta_t^4}{4} \right) \right]. \tag{107} \]

Since the MT branch cut is never crossed, we can take \( \sigma_t = 1 \). Then, summing over both branches of \( \tau(Q_t) \) (see (94)) and choosing

\[ \alpha_t(0,0) = \frac{\sqrt{-2r_t^2}}{2i} \exp(ik_0 t_2) \]  

to satisfy the initial condition (89) ultimately yields

\[ \psi_t[\varepsilon_{\rho}, Q_{t}(\tau)] = \sqrt{1 - 4q_t} \exp(ik_0 q_t) \]

\[ \times \left[ Ai \left( -\frac{r_t^2 \vartheta_t^4}{4} \right) \cos \varpi(t_1) \right. \]

\[ -Bi \left[ -\frac{r_t^2 \vartheta_t^4}{4} \right] \sin \varpi(t_1) \right] \tag{109} \]

where we have chosen \( c = Ai(0) \) and defined

\[ \varpi(t_1) = \frac{2}{3} r_t^3 \vartheta_t^4 - \frac{2}{3} (-4)^{3/2}. \tag{109b} \]

The MGO solution (109) is plotted in figure 1. Notably, this solution is finite along the caustic surface (cutoff) located at \( q_t = 0 \), and agrees remarkably well with the exact solution of (88),

\[ \psi_\infty(q_t) = Ai(q_t) \exp(ik_0 q_t) \]  

A similar plot of (111) is not presented because it is virtually indistinguishable from figure 1.

### 4.3. Imperfectly focused plane wave in uniform medium: cusp caustic

As a final example, let us consider a 2D plane wave described by the paraxial wave equation [31]

\[ i \partial_t \psi(q_t) + \frac{1}{2} \partial_q^2 \psi(q_t) + \psi(q_t) = 0, \tag{112} \]
Fold caustic

$q_2$

$q_1$

**Figure 1.** Contour plot showing the real part of the MGO solution (109), with $k_0 = 2$, near the fold caustic (cutoff) located at $q_3 = 0$ (magenta). The ray trajectories ($q_1(\tau), q_2(\tau)$) are shown as black curves. Note that the field remains finite along the caustic. In fact, the MGO solution is nearly indistinguishable from the exact solution (111).

where $q_1$ is aligned with the optical axis and $q_2$ is transverse to it. Let us assume the initial condition

$$\psi(0, q_2) = \frac{2\pi i}{f} \exp \left(-\frac{i}{2f} q_2^2 - \frac{i a}{4f} q_2^4\right).$$

(113)

This corresponds to a wave that is being focused by an imperfect lens, with focal distance $f$ and aberration $a$. In the absence of aberration ($a = 0$), the initial field will focus at $q = (f, 0)$; however, as shown in figure 2, finite aberration causes the focusing to become distorted, resulting in a cusped wavefield. For simplicity, we shall assume that $f \gg 1$ and $a < 0$.

Equation (112) can be equivalently written as an integral equation (1) with integration kernel

$$D(q, q') = i\partial_q \delta(q - q') - \frac{1}{2} \partial^2_{q'} \delta(q - q') - \delta(q - q').$$

The corresponding Weyl symbol (2) is

$$D(q, k) = k_1 + k_2^2 - 1,$$

(115)

and the corresponding ray equations are

$$\partial_\tau k_1(\tau) = 0, \quad \partial_\tau q_1(\tau) = 1,$n

(116)

Similar to the previous example, let us define $\tau_1$ and $\tau_2$ such that $q_1(0, \tau_2) = 0$ and $q_2(0, \tau_2) = f \tau_2$. Then, imposing the local dispersion relation (4) and the initial condition (113) yields the ray trajectories

$$q_1(\tau) = \tau_1, \quad q_2(\tau) = f \tau_2 + k_2(\tau) \tau_1,$n

(17)

$$k_1(\tau) = 1 - \frac{k_2^2(\tau)}{2}, \quad k_2(\tau) = -\tau_2 - a \tau_2^3.$$

The inverse function $\tau(q)$ is either single- or triple-valued, depending on the value of the discriminant

$$\Delta(q) = 4 \left(\frac{q_1 - f}{aq_1}\right)^3 + 27 \left(\frac{q_2}{aq_1}\right)^2.$$

(118)

If $\Delta(q) > 0$, there is only one ray given by

$$\tau_2^{(0)}(q) = -\sqrt[3]{-\frac{q_2}{2aq_1} + \frac{\Delta(q)}{108}} + \sqrt{\frac{q_2}{2aq_1} - \frac{\Delta(q)}{108}}.$$

(119)

while if $\Delta(q) \leq 0$, there are two additional rays given by

$$\tau_2^{(\pm)}(q) = \text{Re} \left\{ (-1 \pm i\sqrt{3}) \sqrt{-\frac{q_2}{2aq_1} + i \sqrt{\frac{\Delta(q)}{108}}} \right\}.$$

(120)

For all values of $\Delta(q)$, one has $\tau_1(q) = q_1$.

A basis for $Q_1$-space is provided by the vector pair

$$\partial_{\tau_1} z(t) = (1 \ k_2(t) \ 0 \ 0)^T,$$

(121a)

$$\partial_{\tau_2} z(t) = (0 \ j_t \ -k_2(t)k_2(t) \ k_2(t))^T,$$

(121b)

where we have defined

$$j_t = \det \partial_{\tau_2} q(t) = f + t_k^2(t).$$

(122)

Then, $S_t$ is constructed using symplectic Gram–Schmidt orthogonalization [15]. This yields the submatrices

$$A_t = D_t = \frac{1}{\partial_t \varphi_t} \begin{pmatrix} \varphi_t & \varphi_k^2(t) \\ -j_t k_2^2(t) & j_t \end{pmatrix},$$

(123a)

$$B_t = -C_t = \frac{\partial_t k_2^2(t)}{\varphi_t} \begin{pmatrix} 0 & 0 \\ -k_2(t) & 1 \end{pmatrix},$$

(123b)

where we have defined

$$\varphi_t = \sqrt{1 + k_2^2(t)}, \quad \varphi_t = \sqrt{k_2^2(t) + k_2^2(t) \partial_1^2}.$$

(124)

The rotated rays are computed using (8a) and (8b), although the result is quite lengthy and will not be shown. Next, using (27) we compute the envelope as

$$\Phi_t(\tau) = \frac{\varphi_t}{\sqrt{j_t \tau + [1 + k_2(t)k_2(t)]k_2^2(\tau)k_2^2(\tau)}},$$

(125)

where $\Phi_t(\tau) \equiv \Phi_t(Q_t(\tau))$ and we have chosen

$$\Phi_t(Q_t(\tau_1, \tau_2)) = 1.$$n

(126)

Although $\tau(Q_t)$ is impractical to construct explicitly, $\partial_0 \Phi_t$ can still be calculated from $\Phi_t(\tau)$ using

$$\partial_0 \Phi_t(Q_t(t)) = [\partial_0 Q_t(t)]^{-1} \partial_0 \Phi_t(t).$$

(127)

Ultimately, (31) and (32) yield

$$\alpha_t = \alpha_{(0,t)} \sqrt{\frac{\varphi_t(0,t)}{\varphi_t}} \exp \left\{ it_1 \sqrt{1 + \frac{\varphi_t^2}{2}} + \frac{i}{2} \beta_t^2[\varphi_t(t)] \right\},$$

(128)
Thus, \( \alpha_{(0,t_2)} \) is determined by the initial conditions and \( \beta_{t_0}^\rho[q(t)] \) is defined below in (130a).

We next perform an SVD of \( B_t \) to obtain

\[
L = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad R = \frac{1}{\sqrt{t}} \begin{pmatrix} k_2(t) & 1 \\ -1 & k_2(t) \end{pmatrix}.
\] (129)

Hence, we can compute

\[
\beta_{t_0}^\rho[q(t)] = \frac{f^2 c_1^2}{\sqrt{t}} k_2(t) - \frac{2f^2 c_1^2}{\sqrt{t}} k_1(t) + j t_0^2 Q_{1,2}(t),
\] (130a)

\[
\gamma_{t_0}^\rho[q(t)] = \frac{j t_0^2}{2} k_2(t) - K_{1,2}(t) \epsilon_{t_0}.
\] (130b)

Thus, (68) yields

\[
\psi_t[q(t)] = \frac{\sigma_t \alpha_{(0,t_2)} \sqrt{\varphi_{0,t_2}(t)}}{\sqrt{2\pi} k_2(t)} Y_{t_0}^\rho[q(t)] \times \exp \left\{ i t_0 \frac{1 + \sigma_t^2}{2} - i \frac{\beta_{0,t_2}^\rho[q(0,t_2)]}{2} \right\}.
\] (131)

Note that \( k_2(t) \) can change sign, meaning \( \sigma_t \neq 1 \). However, we do not need to explicitly compute \( \sigma_t \) since it will be removed by matching to initial conditions.

After making a slow-envelope approximation, a quartic polynomial (normal form) can be fit to an implicit Taylor expansion of \( \Theta_t[Q_{1,2}(t), Q_{1,2}(t) - \epsilon_{t_0}] \) to ultimately yield

\[
Y_{t_0}^\rho[q(t)] \approx \frac{j t_0^2}{\sqrt{t}} \left( \frac{\sqrt{-2k_2(t)}}{t_1} \right) \exp \left( i t_0 \frac{1 + \sigma_t^2}{2} - i \frac{\beta_{0,t_2}^\rho[q(0,t_2)]}{2} \right) \times \frac{1}{c_{t_2}^2} \int \frac{d \epsilon}{c_{t_2}} \exp \left( i \epsilon \left( i \epsilon + i x \epsilon^2 + i \epsilon^4 \right) \right),
\] (132)

where we have defined

\[
\frac{x_t}{\sqrt{a}} = \left( \frac{f - q_1(t)}{q_1(t)} \right)^{1/4}, \quad y_t = \frac{4f^3}{a} \frac{q_2(t)}{q_1(t)}.
\] (133)

and \( C_t \) is the steepest-descent contour through the saddlepoint \( \epsilon = -t_2 \sqrt{\frac{a}{2}} \).

Our assumption \( a < 0 \) implies that \( \tau(q) \) is single-valued along the initial surface. Thus, (131) yields

\[
\psi_t[q(0,t_2)] = \sigma_t \alpha_{(0,t_2)} \frac{\sqrt{\varphi_{0,t_2}(t)}}{f} \times \exp \left\{ i \frac{1 - \alpha_t^2}{2} - i \frac{\beta_{0,t_2}^\rho[q(0,t_2)]}{2} \right\},
\] (134)

where we have defined \( s_t = \text{sgn}[k_2(t)] \) and have evaluated (132) in the GO limit, since \( f \gg 1 \) implies that the initial surface lies sufficiently far from the caustic. Thus, the initial conditions are satisfied by the choice

\[
\alpha_{(0,t_2)} = \frac{\sqrt{2\pi} \sigma_t}{\sqrt{4\pi t_2}} \exp \left\{ i \frac{\beta_{0,t_2}^\rho[q(0,t_2)]}{2} - i \frac{\beta_{0,t_2}^\rho[q(0,t_2)]}{2} \right\} - \frac{ia}{4} f_2^2 + \frac{2a - 1}{4} \pi.
\] (135)

Equation (38) therefore yields

\[
\psi(q) = \frac{4f}{|a q_1|^{1/4}} \exp \left( i q_1 + i \frac{q_2^2}{2q_1} \right) \times \sum_{\omega \in \tau(q)} \int \frac{d \epsilon}{C_2} \exp \left( i \epsilon q_1 + i \epsilon q_1 q_2^2 + i \epsilon^4 \right),
\] (136)

where the sum is over all real saddlepoint contributions.

Let us recall the Pearcey function [32], defined as

\[
\text{Pe}(x,y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} ds \exp \left( i y s + i x s^2 + i s^4 \right).
\] (137)

Then, when \( \Delta(q) \Delta(q) \geq 0 \), where

\[
\Delta(q) \geq 2 \left( \frac{f - q_1}{|a q_1|} \right)^3 + 27(5 - \sqrt{27}) \left( \frac{q_2^2}{|a q_1|^2} \right)^2
\] (138)

the summation in (136) is simplified as

\[
\sum_{\omega \in \tau(q)} \int \frac{d \epsilon}{C_2} \exp \left( i \epsilon q_1 + i \epsilon q_1 q_2^2 + i \epsilon^4 \right) = \text{Pe}(x_t, y_t).
\] (139)

When \( \Delta(q) \Delta(q) < 0 \) (the caustic shadow), \( \text{Pe}(x,y) \) contains an additional contribution from one of the two complex saddlepoints [33], which is not included in (136). It does not seem possible to isolate the real saddlepoint contribution to \( \text{Pe}(x,y) \) using a complex rotation as done in reference [15] for \( \text{Ai}(x) \). Nevertheless, the shadow contribution is asymptotically subdominant, so within the MGO accuracy we can include it such.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{ray_trajectories.png}
\caption{Ray trajectories for the paraxial equation (112) with initial conditions given by (113). For no aberration (\( a = 0 \)), the rays focus at \( q = (f,0) \). Positive aberration (\( a > 0 \)) causes the outer rays to focus before the focal point, while negative aberration (\( a < 0 \)) causes the outer rays to focus beyond the focal point. Both cases result in a cusped ray pattern.}
\end{figure}
In future work we shall also explore how to include polarization dynamics for modeling vector waves with MGO, including spin–orbit coupling [21, 34–42], mode conversion [21, 43–45], and optical anisotropy [36, 46–48].

Acknowledgments

The work was supported by the US DOE through Contract No. DE-AC02-09CH11466.

Appendix A Symplectic matrices in the SVD basis

As is well-known, the symplectic criterion (15) for $\mathbf{S}$ implies that $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ satisfy [14, 49]

\begin{align}
\mathbf{A}^\top \mathbf{B}^\top - \mathbf{B}^\top \mathbf{A}^\top & = 0_N, \\
\mathbf{B}^\top \mathbf{D} - \mathbf{D}^\top \mathbf{B} & = 0_N, \\
\mathbf{A}^\top \mathbf{D} - \mathbf{C}^\top \mathbf{B} & = 0_N, \\
\mathbf{A}^\top \mathbf{D} - \mathbf{C}^\top \mathbf{B} & = 0_N, \\
\mathbf{C}^\top \mathbf{A} - \mathbf{A}^\top \mathbf{C} & = 0_N, \\
\mathbf{D}^\top \mathbf{C} - \mathbf{D}^\top \mathbf{C} & = 0_N. 
\end{align}

The orthogonality of $\mathbf{L}$ and $\mathbf{R}$ means that $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$, $\tilde{\mathbf{C}}$, and $\tilde{\mathbf{D}}$ also satisfy (A1a)–(A1f). After using the parameterization

\begin{align}
\tilde{\mathbf{A}} &= \begin{pmatrix} a_{\rho\rho} & a_{\rho\varsigma} \\ a_{\varsigma\rho} & a_{\varsigma\varsigma} \end{pmatrix}, \\
\tilde{\mathbf{B}} &= \begin{pmatrix} b_{\rho\rho} & b_{\rho\varsigma} \\ b_{\varsigma\rho} & b_{\varsigma\varsigma} \end{pmatrix}, \\
\tilde{\mathbf{C}} &= \begin{pmatrix} c_{\rho\rho} & c_{\rho\varsigma} \\ c_{\varsigma\rho} & c_{\varsigma\varsigma} \end{pmatrix}, \\
\tilde{\mathbf{D}} &= \begin{pmatrix} d_{\rho\rho} & d_{\rho\varsigma} \\ d_{\varsigma\rho} & d_{\varsigma\varsigma} \end{pmatrix},
\end{align}

(15) becomes

\begin{align}
\begin{pmatrix} (\Lambda_{\rho\rho} - (\Lambda_{\rho\rho})^\top) & (\Lambda_{\rho\varsigma})^\top \\ 0 & 0 \end{pmatrix} & = 0_N. \\
\begin{pmatrix} (\Lambda_{\rho\rho})^\top & (\Lambda_{\rho\varsigma}) \end{pmatrix} & = 0_N. \\
\end{align}

This yields analogous constraints on $\mathbf{D}$, namely

\begin{align}
d_{\rho\rho} = 0_{\rho\rho}, \\
\Lambda_{\rho\rho}d_{\rho\rho} = (\Lambda_{\rho\rho}d_{\rho\rho})^\top.
\end{align}

5. Conclusion

MGO has recently been developed to accurately model caustics by integrating field equations over GO rays. However, as originally formulated in [15], MGO fails to describe what we call quasi-uniform ray patterns (det $\mathbf{B}_k = 0$ over a finite domain). Here, we extend MGO so that the new theory can be applied to any ray pattern in both uniform and non-uniform media. To aid practical implementations, we provide two equivalent representations of the new MGO using either SVDs or Gaussian states.

We demonstrate MGO analytically in three examples, namely, a plane wave propagating in uniform media (no caustic), a plane wave incident on an isolated cutoff (Airy-type Fold caustic), and an imperfectly focused plane wave in a vacuum (Pearcey-type cusp caustic). In all examples, MGO provides an accurate representation of the exact solution that remains finite at the caustics, unlike traditional GO. Yet, as the final example shows, MGO does not always accurately model caustic shadows. Further extensions of MGO to incorporate caustic shadow fields will be investigated in future research.

\[ \psi(q) = \left( \frac{4f}{aq_1^2} \right)^{1/4} \exp \left( iq_1 + i \frac{q_2^2}{2q_1} \right) \]
Next, after using the parameterization 
\[ \tilde{C} = \begin{pmatrix} \mathbf{C}_{pp} & \mathbf{C}_{pc} \\ \mathbf{C}_{cp} & \mathbf{C}_{cc} \end{pmatrix}, \]  
(A8) and used the unitarity of \( \mathbf{L} \) and \( \mathbf{R} \). Next, we must approximate the matrix inverse term to leading order in \( \varepsilon \). Let us adopt the parameterization 
\[ (\hat{\mathbf{B}} + \varepsilon \hat{\mathbf{A}})^{-1} = \begin{pmatrix} \mathbf{m}_{pp} & \mathbf{m}_{pc} \\ \mathbf{m}_{cp} & \mathbf{m}_{cc} \end{pmatrix}. \]  
(B3)

Then, since 
\[ (\mathbf{m}_{pp}(\mathbf{L}_{pp} + \varepsilon \mathbf{a}_{pp}))^{-1} (\mathbf{m}_{cp}(\mathbf{L}_{pp} + \varepsilon \mathbf{a}_{pp}))^{-1} = \mathbf{I}_N, \]  
(B4)

we require 
\[ \varepsilon(\mathbf{m}_{pp}\mathbf{a}_{pp} + \mathbf{m}_{pc}\mathbf{a}_{cc}) = 0, \]  
(B5a)
\[ \varepsilon(\mathbf{m}_{cp}\mathbf{a}_{pp} + \mathbf{m}_{cc}\mathbf{a}_{cc}) = 0, \]  
(B5b)
\[ \mathbf{m}_{pp}(\mathbf{L}_{pp} + \varepsilon \mathbf{a}_{pp}) = \mathbf{I}_N, \]  
(B5c)
\[ \mathbf{m}_{pp}(\mathbf{L}_{pp} + \varepsilon \mathbf{a}_{pp}) = \mathbf{I}_N. \]  
(B5d)

Solving (B5a)–(B5d) in sequence yields
\[ \mathbf{m}_{cc} = \varepsilon^{-1} \mathbf{a}_{cc}^{-1}, \]  
(B6)
where we have used the fact that \( \mathbf{a}_{cc} \) is invertible. Hence, we obtain the following leading-order approximations:
\[ (\hat{\mathbf{D}} + \varepsilon \hat{\mathbf{C}})(\hat{\mathbf{B}} + \varepsilon \hat{\mathbf{A}})^{-1} \approx \begin{pmatrix} \mathbf{a}_{pp}^{-1} & 0 \\ 0 & \mathbf{a}_{cc}^{-1} \end{pmatrix}, \]  
(B7)
\[ (\hat{\mathbf{B}} + \varepsilon \hat{\mathbf{A}})^{-1} \approx \begin{pmatrix} \mathbf{L}_{pp}^{-1} & 0 \\ 0 & \varepsilon^{-1} \mathbf{L}_{cc}^{-1} \end{pmatrix}. \]  
(B8)

Upon introducing the vector decompositions
\[ \mathbf{R}^\top \mathbf{x} = \begin{pmatrix} \mathbf{x}_p \\ \mathbf{x}_c \end{pmatrix}, \quad \mathbf{L}^\top \mathbf{y} = \begin{pmatrix} \mathbf{y}_p \\ \mathbf{y}_c \end{pmatrix} \]  
(B9) where all subvectors \( \nu_\nu \) are size \( n \times 1 \), we obtain
\[ \tilde{\mathbf{G}}(\mathbf{q}, \mathbf{Q}) \approx \begin{pmatrix} \mathbf{q}_c - \mathbf{a}_{cc}^{-1} \mathbf{Q}_c + \frac{\mathbf{q}_c^\top \Lambda_{cc}^{-1} \mathbf{a}_{cc}}{2} + \frac{\mathbf{q}_c^\top \mathbf{Q}_c \mathbf{M}_1 \mathbf{L}^\top \mathbf{Q} + \frac{\mathbf{q}_c^\top \mathbf{Q}_c \mathbf{M}_2 \mathbf{L}^\top \mathbf{Q}}{2} \end{pmatrix} \]  
(B2) where the matrices \( \mathbf{M}_1 \) and \( \mathbf{M}_2 \) are defined in (51a) and (51b).
Appendix C MTs in the mixed basis of configuration and coherent states

Here, we derive (72) using Gaussian coherent states. To help with the presentation, in this section, we employ the bra-ket notation of quantum mechanics [50].

The Gaussian coherent states, denoted |Z₀⟩, are defined by their Q-space representations

\[ \langle Q|Z₀ \rangle = \frac{\exp \left[ -\frac{\mathbf{Q} \cdot Q}{2} + i \mathbf{K}_0 \cdot (\mathbf{Q} - \frac{Q}{2}) \right]}{\pi^{N/4}}, \]  
(C1)

where |Q⟩ are the eigenstates of the Q-space position operator ˆQ, normalized as 〈Q|Q′⟩ = δ(Q − Q′). The parameters Q₀ and K₀ define the center of |Z₀⟩ in phase space, that is,

\[ \langle Z₀|Q|Z₀ \rangle = Q₀, \quad \langle Z₀|P|Z₀ \rangle = K₀, \]  
(C2)

where ˆP is the Q-space momentum operator. The states |Z₀⟩ also satisfy a completeness relation of the form

\[ \hat{1} = \int \frac{dQ₀dK₀}{(2\pi)^N} |Z₀⟩⟨Z₀|, \]  
(C3)

where 1 denotes the identity operator.

Let us denote ψ₁(q) by |ψ₁⟩ and ψ₁(Q) by ⟨Q|ψ₁⟩. Then, the inverse MT of (34) can be written as

\[ ψ₁(q) = \int dQ ⟨Q|q⟩⟨Q|ψ₁⟩ = \int dQ ⟨Q|Q⟩ψ(Q), \]  
(C4)

where we have used the completeness of |Q⟩, that is,

\[ \hat{1} = \int dQ ⟨Q|Q⟩. \]  
(C5)

The matrix element ⟨q|Q⟩ is the inverse MT kernel. Direct computation of ⟨q|Q⟩ will yield (35), which is delta-shaped when detB₀ = 0. However, by introducing the normalizable states |Z₀⟩ as

\[ ⟨q|Q⟩ = \int \frac{dQ₀dK₀}{(2\pi)^N} ⟨q|Z₀⟩⟨Z₀|Q⟩, \]  
(C6)

a nonsingular representation of the MT can be obtained. Indeed, as shown in [23],

\[ ⟨q|Z₀⟩ = \sigma \exp \left[ -\frac{q \cdot [D₀ - iB₀]^{-1} (A₀ + iC₀) q}{\pi^{N/4} \sqrt{\det(D₀ - iB₀)}} \right] \times \exp \left[ \left( q - \frac{i}{2} B₀^T \zeta \right)^T (D₀ - iB₀)^{-1} \zeta - \frac{1}{2} Q₀^T \zeta \right], \]  
(C7)

where we have introduced the complex vector \( \zeta \equiv Q₀ + iK₀ \). Importantly, the complex matrix \( D₀ - iB₀ \) is always invertible [29]. Then, upon using \( ⟨Z₀|Q⟩ = (⟨Q|Z₀⟩)^* \), one computes

\[ \langle q|Z₀⟩⟨Z₀|Q⟩ = \frac{\sigma \exp \left[ -\frac{q \cdot [D₀ - iB₀]^{-1} (A₀ + iC₀) q}{\pi^{N/4} \sqrt{\det(D₀ - iB₀)}} \right] \times \exp \left[ \left( q - \frac{i}{2} B₀^T \zeta \right)^T (D₀ - iB₀)^{-1} \zeta - \frac{1}{2} Q₀^T \zeta \right]}{\pi^{N/2} \sqrt{\det(D₀ - iB₀)}} \times \exp \left[ \left( q - \frac{i}{2} B₀^T \zeta \right)^T (D₀ - iB₀)^{-1} \zeta - \frac{1}{2} Q₀^T \zeta \right] \right]. \]  
(C8)

Finally, we integrate over Q₀ in (C6) to obtain (72). Note that K₀ cannot be integrated over without inverting B₀.

ORCID iDs

N A Lopez https://orcid.org/0000-0001-5566-3731
I Y Dodin https://orcid.org/0000-0003-0243-6257

References

[1] Kravtsov Y A and Orlov Y I 1990 Geometrical Optics of Inhomogeneous Media (Berlin: Springer)
[2] Tracy E R, Brizard A J, Richardson A S and Kaufman A N 2014 Ray Tracing and Beyond: Phase Space Methods in Plasma Wave Theory (Cambridge: Cambridge University Press) (https://doi.org/10.1017/CBO9780511667565)
[3] Kravtsov Y A and Orlov Y I 1993 Caustics, Catastrophes and Wave Fields (Berlin: Springer) (https://doi.org/10.1007/978-3-642-59887-6)
[4] Berry M V and Upstill C 1980 Catastrophe optics: morphologies of caustics and their diffraction patterns. Prog. Opt. 18 257
[5] Berry M V and Upstill C 1980 Catastrophe optics: morphologies of caustics and their diffraction patterns. Prog. Opt. 18 257
[6] Hobbs C A, Connor J N L and Kirk N P 2007 Theory and numerical evaluation of oddoids and evenoids: oscillatory cuspid integrals with odd and even polynomial phase functions J. Comput. Appl. Math. 207 192
[7] Borghi R 2016 Computational optics through sequence transformations Prog. Opt. 61 1
[8] Zannotti A, Diebel F, Boguslawski M and Denz C 2017 Optical catastrophes of the swallowtail and butterfly beams New J. Phys. 19 053004
[9] Espindola-Ramos E et al 2019 Paraxial optical fields whose intensity pattern skeletons are stable caustics J. Opt. Soc. Am. A 36 1820
[10] Wright J C, Bonoli P T, Schmidt A E, Phillips C K, Valeo E J, Harvey R W and Brambilla M A 2009 An assessment of full wave effects on the propagation and absorption of lower hybrid waves Phys. Plasmas 16 072502
[11] Shiraishi S, Meneghini O, Parker R, Bonoli P, Garrett M, Kaufman M C, Wright J C and Wukitch S 2010 Plasma wave simulation based on a versatile finite element method solver Phys. Plasmas 17 056119
[12] Myatt J F, Follett R K, Shaw J G, Edgell D H, Froula D H, Igumenschev I V and Goncharov V N 2017 A wave-based model for cross-beam energy transfer in direct-drive inertial confinement fusion Phys. Plasmas 24 056308
[13] Maslov V P and Fedoriuk M V 1981 Semiclassical Approximation in Quantum Mechanics (Dordrecht: Reidel)
[14] Lopez N A and Dodin I Y 2019 Pseudo-differential representation of the metaplectic transform and its application to fast algorithms J. Opt. Soc. Am. A 36 1846
[15] Lopez N A and Dodin I Y 2020 Restoring geometrical optics near caustics using sequenced metaplectic transforms New J. Phys. 22 083078
[16] Garcia-Bulle M, Lassner W and Wolf K B 1986 The metaplectic group within the heisenberg-weyl ring J. Math. Phys. 27 29
[17] Bacry H and Cadilhac M 1981 Metaplectic group and fourier optics Phys. Rev. A 23 2533
[18] Sudarshan E C G, Simon R and Mukunda N 1983 Paraxial-wave optics and relativistic front description. i. the scalar theory Phys. Rev. A 28 2921
[19] Mukunda N, Simon R and Sudarshan E C G 1983 Paraxial-wave optics and relativistic front description. ii. the vector theory Phys. Rev. A 28 2933
[20] Arnold V I 1975 Critical points of smooth functions and their normal forms Russ. Math. Surv. 30 1
[21] Dodin I Y, Ruiz D E, Yanagihara K, Zhou Y and Kubo S 2019 Quasioptical modeling of wave beams with and without mode conversion: I. basic theory Phys. Plasmas 26 072110
[22] Dodin I Y, Zhmoginov A I and Ruiz D E 2017 Variational principles for dissipative (sub)systems, with applications to the theory of linear dispersion and geometrical optics Phys. Lett. A 381 1411
[23] Littlejohn R G 1986 The semiclassical evolution of wave packets Phys. Rep. 138 193
[24] Lohmann A W 1993 Image rotation, Wigner rotation and the fractional Fourier transform J. Opt. Soc. Am. A 10 2181
[25] Heller E J 1977 Generalized theory of semiclassical amplitudes J. Chem. Phys. 66 5777
[26] Chester C, Friedman B and Ursell F 1957 An extension of the method of steepest descents Proc. Cambridge Philos. Soc. 53 599
[27] Bleistein N and Handelsman R A 1986 Asymptotic Expansions of Integrals (New York: Dover)
[28] Scully M O and Zubairy M S 2012 Quantum Optics (Cambridge: Cambridge University Press)
[29] Littlejohn R G and Robbins J M 1987 New way to compute Maslov indices Phys. Rev. A 36 2953
[30] Olver F W J, Lozier D W, Boisvert R F and Clark C W 2010 NIST Handbook of Mathematical Functions (Cambridge: Cambridge University Press)
[31] Kogelnik H and Li T 1966 Laser beams and resonators Appl. Opt. 5 1550
[32] Paris R B 1991 The asymptotic behaviour of Peccey's integral for complex variables Proc. R. Soc. A 432 391
[33] Wright F J 1980 The stokes set of the cusp diffraction catastrophe J. Phys. A: Math. Gen. 13 2913
[34] Bliokh K Y and Bliokh Y P 2004 Modified geometrical optics of a smoothly inhomogeneous isotropic medium: the anisotropy, Berry phase and the optical magnus effect Phys. Rev. E 70 026605
[35] Bliokh K Y and Bliokh Y P 2004 Topological spin transport of photons: the optical magnus effect and berry phase Phys. Lett. A 333 181
[36] Ruiz D E and Dodin I Y 2017 Extending geometrical optics: a Lagrangian theory for vector waves Phys. Plasmas 24 055504
[37] Ruiz D E and Dodin I Y 2015 First-principles variational formulation of polarization effects in geometrical optics Phys. Rev. A 92 043805
[38] Ruiz D E and Dodin I Y 2015 Lagrangian geometrical optics of nonadiabatic vector waves and spin particles Phys. Lett. A 379 2337
[39] Ruiz D E, Ellison C L and Dodin I Y 2015 Relativistic ponderomotive Hamiltonian of a Dirac particle in a vacuum laser field Phys. Rev. A 92 062124
[40] Ruiz D E 2017 A geometric theory of waves and its applications to plasma physics PhD Thesis Princeton University (https://arxiv.org/abs/1708.05423)
[41] Oancea M A, Joudioux J, Dodin I Y, Ruiz D E, Paganini C F and Andersson L 2020 The gravitational spin Hall effect of light Phys. Rev. D 102 024075
[42] Bliokh K Y, Rodriguez-Fortuno F J, Nori F and Zayats A V 2015 Spin-orbit interactions of light Nat. Photon. 9 796
[43] Yanagihara K, Dodin I Y and Kubo S 2019 Quasioptical modeling of wave beams with and without mode conversion: II. Numerical simulations of single-mode beams Phys. Plasmas 26 072111
[44] Yanagihara K, Dodin I Y and Kubo S 2019 Quasioptical modeling of wave beams with and without mode conversion: III. Numerical simulations of mode-converting beams Phys. Plasmas 26 072112
[45] Dodin I Y, Ruiz D E and Kubo S 2017 Mode conversion in cold low-density plasma with a sheared magnetic field Phys. Plasmas 24 122116
[46] Kravtsov Y A, Naida O N and Fuki A A 1996 Waves in weakly anisotropic 3D inhomogeneous media: quasi-isotropic approximation of geometrical optics Phys.-Usp. 39 129
[47] Kravtsov Y A, Bieg B and Bliokh K Y 2007 Stokes-vector evolution in a weakly anisotropic inhomogeneous medium J. Opt. Soc. Am. A 24 3388
[48] Bliokh K Y, Frolov D Y and Kravtsov Y A 2007 Non-Abelian evolution of electromagnetic waves in a weakly anisotropic inhomogeneous medium Phys. Rev. A 75 053821
[49] Luneburg R K 1964 Mathematical Theory of Optics (Berkeley, CA: University of California Press)
[50] Stoler D 1981 Operator methods in physical optics J. Opt. Soc. Am. 71 334