MULTISCALE MODEL. SIMUL.
Vol. 13, No. 3, pp. 733–765

A GAUSSIAN BEAM METHOD FOR HIGH FREQUENCY SOLUTION OF SYMMETRIC HYPERBOLIC SYSTEMS WITH POLARIZED WAVES*

LELAND JEFFERIS† AND SHI JIN‡

Abstract. Symmetric hyperbolic systems include many physically relevant systems of PDEs such as Maxwell's equations, the elastic wave equations, and the acoustic equations [L. Ryzhik, G. Papanicolaou, and J. Keller, *Wave Motion*, 24 (1996), pp. 327–370]. In the current paper we extend the Gaussian beam method to efficiently compute the high frequency solutions to such systems with polarized waves, in which the dispersion matrix of the hyperbolic system has eigenvalues with constant multiplicity greater than one over the domain of computation. The new results in this paper include new Gaussian beam equations in the presence of multiple eigenvalues, error estimates for Gaussian beam summation in the symmetric hyperbolic system case, and a new multidirectional Eulerian summation formula which maintains accuracy after the formation of caustics.

Key words. Gaussian beam, high frequency waves, polarized waves

AMS subject classifications. 00A69, 74J05

DOI. 10.1137/130935318

1. Introduction.

We will study the general symmetric hyperbolic system of the form[1]

$$\begin{cases} A(\mathbf{x})\frac{\partial \mathbf{u}_\varepsilon}{\partial t} + D^j \frac{\partial \mathbf{u}_\varepsilon}{\partial x_j} = 0, \\ \mathbf{u}_\varepsilon(\mathbf{x},0) = \mathbf{u}_0(\mathbf{x}) e^{\mathbf{i}S_0(\mathbf{x})/\varepsilon}, \end{cases} \tag{1.1}$$

where $\mathbf{u} \in \mathbb{C}^n$, $\mathbf{x} \in \mathbb{R}^d$, $A(\mathbf{x})$ is symmetric positive definite, and the D^j are symmetric and independent of $\mathbf{x}$ and t. Many physical problems such as Maxwell's equations, the elastic wave equations, and the acoustic equations all may be written in such a form with the correct choices of $A(\mathbf{x})$ and the D^j, and these particular three examples were discussed in [28]. In many physical applications, ε, which characterizes the wave length, is very small compared to the scale of the computational domain, and the numerical meshes and time steps need to resolve this small scale; thus computing the high frequency solutions, particularly in high dimensions, is prohibitively expensive.

One efficient way to deal with high frequency wave problems is to solve the limiting equation by finding the asymptotic equation when $\varepsilon \to 0$. The Wigner transform, introduced in [34], is a powerful mathematical tool for studying this limit [8, 20, 28], since it is valid globally in time, even beyond caustic formation. The limiting equation is the Liouville equation which does not depend on ε, permitting large time steps and

*Received by the editors September 3, 2013; accepted for publication (in revised form) January 21, 2015; published electronically July 2, 2015. This work was partially supported by NSF grants DMS-1114546 and DMS-1107291: NSF Research Network in Mathematical Sciences "KI-Net: Kinetic description of emerging challenges in multiscale problems of natural sciences."
http://www.siam.org/journals/mms/13-3/93531.html

†Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706 (jefferis.l@gmail.com).

‡Department of Mathematics, Institute of Natural Sciences and Ministry of Education Key Lab of Scientific and Engineering Computing, Shanghai Jiao Tong University, Shanghai 20040, China, and Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706 (jin@math.wisc.edu).

[1]Assume the following summation convention: Repeated Latin indices are summed, whereas repeated Greek indices are not summed.

mesh sizes. In a previous paper [11], we developed a numerical method based on this approach for the problem under study. One should note that the Liouville equation–based classical or geometric optic limit approach, derived via the limit $\varepsilon \to 0$, does not offer good accuracy near the caustics.

A more accurate approach is the Gaussian beam method, originally introduced in [10, 25]. The key idea in all Gaussian beam methods is to take advantage of the fact that high frequency waves have particle-like properties. Specifically, one decomposes the initial wave function into localized wave packets (Gaussian beams) which are then evolved individually along particle trajectories and finally summed up to construct the solution at a later time. It was first studied rigorously in [27] and has seen many recent developments in both Eulerian and Lagrangian frameworks [14, 15, 16, 19, 22, 23], error estimates [2, 21], and fast Gaussian wave decompositions [1, 31, 26]. A related approach, known as the Hagedorn wave packet method, was studied in [9, 7]. For recent surveys for semiclassical computational methods for high frequency waves, see [6, 12].

The main difficulties for Gaussian beam methods in the nonstrict hyperbolic systems are threefold. First, the multiple eigenvalues of the dispersion matrix combined with the complex valued phase introduced in the Gaussian beam ansatz lead to analytic difficulties which must be delicately dealt with. Second, one must determine a meaningful ansatz for Gaussian beam solutions which are to represent the polarized waves corresponding to multiple eigenvalues of the dispersion matrix. Third, one must derive a meaningful coupling matrix (to represent polarized wave propagation) with the correct analytic properties. In this paper, all of these difficulties have been resolved. To handle the complex phase of the Gaussian beam ansatz, we introduce a notion of *spectrum preserving* which any system must satisfy in order for our Gaussian beam method to be applicable. We show, however, that many physically relevant equations all satisfy this definition and are thus solvable by our method. The second difficulty was overcome simply by making a careful choice and, by trial and error, showing that our chosen ansatz gives meaningful results. To overcome the last difficulty, we derive a form for our coupling matrix which matches the one discovered in [28], thereby showing a deep connection between this new Gaussian beam method and our previous work [11].

In addition to developing a Gaussian beam method for nonstrict symmetric hyperbolic systems, we also obtain further supplementary results. Convergence results for the decomposition of the initial condition are presented in [21, 30]; here we extend and modify these convergence results to suit our case. The final summing process for Eulerian Gaussian beams was shown to lose accuracy after the formation of caustics in [14]; here we introduce a new Eulerian summation formula to solve this problem. Numerical results are also provided in one and two dimensions to demonstrate the effectiveness of our method; these results pave the way toward heftier three-dimensional simulations which could be well handled by our provided method.

We also would like to point out that some of the difficulties faced in this system are shared in quantum dynamics with band-crossings. One example is the surface hopping phenomenon in which particles tunnel through different electronic potential surfaces and the classical Bohn–Oppenheimer approximation breaks down [32, 18, 13]. Another difficulty is the crossing of Bloch bands in the Schrödinger equation with periodic potentials [29, 4]. The method developed in this paper sheds light on these important physical and chemical problems.

The paper is outlined as follows. Section 2 introduces the Gaussian beam method in the Lagrangian framework and proves essential results pertaining to this method's

convergence and boundedness. Section 3 introduces the Gaussian beams in the Eulerian framework, introduces a new Eulerian summation formula, and provides further simplifications to the method in the one-dimensional case. Section 4 contains numerical results which include simulations in one and two dimensions as well as convergence tests. Finally, section 5 contains our conclusions.

2. A Gaussian beam method.

2.1. The Lagrangian Gaussian beam method. The Gaussian beam ansatz for (1.1) is[2]

$$\phi_\varepsilon^{la}(\mathbf{x},t) = [\mathbf{a}_0(\mathbf{x},\mathbf{q},t) + \varepsilon \mathbf{a}_1(\mathbf{x},\mathbf{q},t) + \cdots] e^{\frac{\mathbf{i}}{\varepsilon} T(\mathbf{x},\mathbf{q},t)}, \tag{2.1}$$

where

$$T(\mathbf{x},\mathbf{q},t) = S(\mathbf{q},t) + \mathbf{p}(\mathbf{q},t)\cdot(\mathbf{x}-\mathbf{q}) + \frac{1}{2}(\mathbf{x}-\mathbf{q})^T M(\mathbf{q},t)(\mathbf{x}-\mathbf{q}), \tag{2.2}$$

$\mathbf{q} = \mathbf{q}(\mathbf{q}_0,t) \in \mathbb{R}^d$, $\mathbf{p}(\mathbf{q},t) \in \mathbb{R}^d$, and $M \in \mathbb{C}^{d\times d}$ is assumed to be symmetric with a positive definite imaginary part (so that (2.1) has a Gaussian profile). Define

$$\mathcal{J} \equiv A(\mathbf{x}) d_t + D^j \partial_{x_j}, \tag{2.3}$$

where

$$d_t \equiv \frac{\partial}{\partial t} + (\partial_t \mathbf{q})\cdot \nabla_{\mathbf{q}} \tag{2.4}$$

is our notation for the total time derivative, which we distinguish from the symbol $\frac{\partial}{\partial t}$ (appearing in (1.1)) for clarity. Substituting (2.1) into (1.1) and keeping the first two orders in ε, one obtains

$$\begin{aligned} &\mathcal{O}(1/\varepsilon): \quad (\mathcal{J}T)\mathbf{a}_0 = \mathbf{0}, \\ &\mathcal{O}(1): \qquad \mathcal{J}\mathbf{a}_0 + \mathbf{i}(\mathcal{J}T)\mathbf{a}_1 = \mathbf{0}. \end{aligned} \tag{2.5}$$

By first multiplying by A^{-1} on the left side, the $\mathcal{O}(1/\varepsilon)$ equation of (2.5) may be written as

$$[d_t T + A^{-1} D^j \partial_{x_j} T]\mathbf{a}_0 = \mathbf{0}. \tag{2.6}$$

We define the *dispersion matrix* as

$$L(\mathbf{x},\mathbf{k}) \equiv A^{-1}(\mathbf{x}) k_i D^i \tag{2.7}$$

so that we may write (2.6) as

$$[(d_t T)\, I + L(\mathbf{x}, \nabla_{\mathbf{x}} T)]\, \mathbf{a}_0 = \mathbf{0}. \tag{2.8}$$

Let $\langle \mathbf{u}, \mathbf{v}\rangle$ be the standard inner product on $\mathbb{C}^n$; then define the new inner product

$$\langle \mathbf{u}, \mathbf{v}\rangle_A \equiv \langle A\mathbf{u}, \mathbf{v}\rangle \tag{2.9}$$

[2]In this paper, $\mathbf{i}$ denotes the imaginary unit $\sqrt{-1}$ and is not a vector. i, when used, is an index.

under which $L(\mathbf{x},\mathbf{k})$ is self-adjoint when $\mathbf{x}$ and $\mathbf{k}$ are real. Note that the self-adjoint property of $L(\mathbf{x},\mathbf{k})$ guarantees that it has a complete set of eigenvectors when $\mathbf{x}$ and $\mathbf{k}$ are real. We introduce the following definition.

DEFINITION 2.1. *A system of the form* (1.1) *with dispersion matrix* $L(\mathbf{x},\mathbf{k})$ *defined by* (2.7) *is* nonstrict hyperbolic *at some point* $(\mathbf{x},\mathbf{k})$ *if* $L(\mathbf{x},\mathbf{k})$ *has a multiple eigenvalue at that point. The system is* constantly nonstrict hyperbolic *on some domain* $\mathcal{D} \subset \mathbb{R}^{2d}$ *if the multiplicity (geometric and algebraic) of the eigenvalues of* $L(\mathbf{x},\mathbf{k})$ *remains constant over all points* $(\mathbf{x},\mathbf{k}) \in \mathcal{D}$.

For the remainder of this paper we consider solutions on domains of constant nonstrict hyperbolicity as defined in Definition 2.1. Let $H_\tau(\mathbf{x},\mathbf{k})$ be an eigenvalue of $L(\mathbf{x},\mathbf{k})$ with multiplicity r. Note that by our assumption in Definition 2.1, r remains constant within the domain of computation. Here τ enumerates the unique eigenvalues of $H_\tau(\mathbf{x},\mathbf{k})$ or $L(\mathbf{x},\mathbf{k})$. Let $\mathbf{b}^{\tau,s}(\mathbf{x},\mathbf{k})$ for $s = 1,\dots,r$ be the eigenvectors corresponding to $H_\tau(\mathbf{x},\mathbf{k})$ so that

$$\langle \mathbf{b}^{\tau,i}, \mathbf{b}^{\tau,j} \rangle_A = \delta_{ij}. \tag{2.10}$$

The next step is to observe that (2.8) implies that

$$d_t T + H_\tau(\mathbf{x}, \tilde{\mathbf{p}}) = 0 \quad \text{and} \quad \mathbf{a}_0 = \sum_{s=1}^{r} c_s(\mathbf{q},t)\mathbf{b}^{\tau,s}(\mathbf{x},\tilde{\mathbf{p}}), \tag{2.11}$$

where

$$\tilde{\mathbf{p}} \equiv \nabla_{\mathbf{x}} T. \tag{2.12}$$

In (2.11), $\mathbf{a}_0$ spans the eigenspace of H_τ when $\mathbf{x}$ and $\mathbf{k}$ are real. However, it is only when $\mathbf{x}$ and $\mathbf{k}$ are real that $L(\mathbf{x},\mathbf{k})$ is self-adjoint in $\langle\cdot,\cdot\rangle_A$. Since $\tilde{\mathbf{p}} = \nabla_{\mathbf{x}} T$ is complex, the dispersion matrix $L(\mathbf{x},\tilde{\mathbf{p}})$ has no guaranteed nice structure of its spectrum. Thus for (2.11) to be well defined, we must assume that we have an expression for our complex eigenvalues and eigenvectors which is valid at least when $\tilde{\mathbf{p}}$ has a small imaginary part. We formalize this assumption with a definition.

DEFINITION 2.2. *A system* (1.1) *with dispersion matrix* $L(\mathbf{x},\mathbf{k})$ *given by* (2.7) *is* spectrum preserving *on some domain* $\mathcal{D} \subset \mathbb{R}^{2d}$ *if in addition to having constant nonstrict hyperbolicity on* $\mathcal{D}$ *(see Definition* 2.1*), the multiplicity of the eigenvalues (algebraic and geometric) is also preserved when* $|\mathrm{Im}(\mathbf{k})| < \delta$ *for some fixed* $\delta > 0$.

Remark 2.3. We have conjectured that a dispersion matrix $L(\mathbf{x},\mathbf{k})$ which has constant nonstrict hyperbolicity in the sense of Definition 2.1 on some domain $\mathcal{D} \subset \mathbb{R}^{2d}$ is automatically spectrum preserving in the sense of Definition 2.2 on the same domain $\mathcal{D}$. This is easy to show in the $d = 1$ case (it follows easily from the observations about one-dimensional systems presented in section 3.1), but the general $d > 1$ case is not so simple. How to prove or disprove this conjecture remains an open question. Nevertheless, in Appendix A we prove that Definition 2.2 holds in the case of the three-dimensional acoustic equations, Maxwell's equations, and the elastic wave equations, respectively. Definition 2.2 also holds for the model problems considered in section 4, but we omit the relatively simple proofs.

From this point forward we will assume that the system (1.1) is spectrum preserving (Definition 2.2) on the domain of our interest. With this assumption, the expansion shown in (2.11) is now valid. Taking derivatives of the first equation of

(2.11) with respect to $\mathbf{x}$, one obtains

(2.13)

$$\begin{aligned}
&\text{0th: } \partial_t T + (\partial_t \mathbf{q}) \cdot \nabla_{\mathbf{q}} T + H_\tau = 0,\\
&\text{1st: } \partial_t \nabla_{\mathbf{x}} T + (\partial_t \mathbf{q}) \cdot \nabla_{\mathbf{xq}} T + \nabla_{\mathbf{x}} H_\tau + \nabla_{\tilde{\mathbf{p}}} H_\tau \cdot \nabla_{\mathbf{xx}} T = \mathbf{0},\\
&\text{2nd: } \partial_t \nabla_{\mathbf{xx}} T + (\partial_t \mathbf{q}) \cdot \nabla_{\mathbf{xxq}} T + \nabla_{\mathbf{xx}} H_\tau + \nabla_{\mathbf{x}\tilde{\mathbf{p}}} H_\tau \nabla_{\mathbf{xx}} T\\
&\qquad\qquad + \nabla_{\mathbf{xx}} T \nabla_{\tilde{\mathbf{p}}\mathbf{x}} H_\tau + \nabla_{\mathbf{xx}} T \nabla_{\tilde{\mathbf{p}}\tilde{\mathbf{p}}} H_\tau \nabla_{\mathbf{xx}} T + \nabla_{\tilde{\mathbf{p}}} H_\tau \nabla_{\mathbf{xxx}} T = 0I,
\end{aligned}$$

where in the above $H_\tau = H_\tau(\mathbf{q}, \tilde{\mathbf{p}})$. Evaluating (2.13) at $\mathbf{x} = \mathbf{q}$ gives

(2.14)

$$\begin{aligned}
&\partial_t S + (\partial_t \mathbf{q}) \cdot (\nabla_{\mathbf{q}} S - \mathbf{p}) + H_\tau = 0,\\
&\partial_t \mathbf{p} + (\partial_t \mathbf{q}) \cdot (\partial_{\mathbf{q}} p - M) + \nabla_{\mathbf{q}} H_\tau + (\nabla_{\mathbf{p}} H_\tau) M = \mathbf{0},\\
&\partial_t M + (\partial_t \mathbf{q}) \cdot (\nabla_{\mathbf{q}} M) + \nabla_{\mathbf{qq}} H_\tau + (\nabla_{\mathbf{qp}} H_\tau) M + M(\nabla_{\mathbf{pq}} H_\tau) + M(\nabla_{\mathbf{pp}} H_\tau) M = 0I,
\end{aligned}$$

where now $H_\tau = H_\tau(\mathbf{q}, \mathbf{p})$. Finally, set $(\partial_t \mathbf{q}) = \nabla_{\mathbf{p}} H_\tau$ to obtain

$$\begin{aligned}
d_t \mathbf{q} &= \nabla_{\mathbf{p}} H_\tau,\\
d_t \mathbf{p} &= -\nabla_{\mathbf{q}} H_\tau,\\
d_t S &= \nabla_{\mathbf{p}} H_\tau \cdot \mathbf{p} - H_\tau,\\
d_t M &= -\nabla_{\mathbf{qq}} H_\tau - \nabla_{\mathbf{qp}} H_\tau M - M \nabla_{\mathbf{pq}} H_\tau - M \nabla_{\mathbf{pp}} H_\tau M.
\end{aligned} \tag{2.15}$$

The first two equations of (2.15) are the *ray tracing equations* or *bicharacteristic equations* which track the center of the Gaussian beam and form a Hamiltonian system. For now, we assume that the matrix M is symmetric for all time (proved in section 2.2).

The solvability conditions for the $\mathcal{O}(1)$ equation of (2.5) are

$$\langle \bar{\mathbf{b}}^{\tau,i}(\mathbf{x}, \tilde{\mathbf{p}}), \mathcal{J}\mathbf{a}_0 \rangle = 0 \quad \text{for} \quad i = 1, \ldots, r, \tag{2.16}$$

where we note that the conjugate $\bar{\mathbf{b}}^{\tau,i}(\mathbf{x}, \tilde{\mathbf{p}})$ is used in place of $\mathbf{b}^{\tau,i}(\mathbf{x}, \tilde{\mathbf{p}})$ since the adjoint of $\mathcal{J}T$ is not itself but $\overline{\mathcal{J}T}$. Substituting our assumed form for $\mathbf{a}_0$ given in (2.11) into (2.16) gives, after some simplifications,

(2.17)

$$\sum_{s=1}^{r} \left\{ [d_t c_s(\mathbf{q}, t)] \, \langle \bar{\mathbf{b}}^{\tau,i}(\mathbf{x}, \tilde{\mathbf{p}}), \mathbf{b}^{\tau,s}(\mathbf{x}, \tilde{\mathbf{p}}) \rangle_A + c_s(\mathbf{q}, t) \langle \bar{\mathbf{b}}^{\tau,i}(\mathbf{x}, \tilde{\mathbf{p}}), \mathcal{J}\mathbf{b}^{\tau,s}(\mathbf{x}, \tilde{\mathbf{p}}) \rangle \right\} = 0.$$

Next, observe that

$$\begin{aligned}
A d_t \mathbf{b}^{\tau,s} &= [(d_t \mathbf{p})_n - (M d_t \mathbf{q})_n] A \partial_{\tilde{\mathbf{p}}_n} \mathbf{b}^{\tau,s},\\
D^j d_{\mathbf{x}_j} \mathbf{b}^{\tau,s} &= D^j [\partial_{\mathbf{x}_j} \mathbf{b}^{\tau,s} + \{(\nabla_{\tilde{\mathbf{p}}} \mathbf{b}^{\tau,s}) M\} \mathbf{e}^j],
\end{aligned} \tag{2.18}$$

where $\mathbf{e}^j$ is the jth coordinate vector. Letting $\mathbf{x} = \mathbf{q}$, using (2.15), and rearranging terms gives that

$$\mathcal{J}\mathbf{b}^{\tau,s}|_{\mathbf{x}=\mathbf{q}} = D^j \left\{ \nabla_{\mathbf{q}} \mathbf{b}^{\tau,s} + (\nabla_{\mathbf{p}} \mathbf{b}^{\tau,s}) M \right\} \mathbf{e}^j - A \nabla_{\mathbf{p}} \mathbf{b}^{\tau,s} [\nabla_{\mathbf{q}} H_\tau + M \nabla_{\mathbf{p}} H_\tau]. \tag{2.19}$$

Define the matrix E^τ as

$$E^\tau_{is} = \langle A\mathbf{b}^{\tau,i}, A^{-1}D^j\{\nabla_{\mathbf{q}}\mathbf{b}^{\tau,s} + (\nabla_{\mathbf{p}}\mathbf{b}^{\tau,s})M\}\mathbf{e}^j - \nabla_{\mathbf{p}}\mathbf{b}^{\tau,s}[\nabla_{\mathbf{q}}H_\tau + M\nabla_{\mathbf{p}}H_\tau]\rangle. \tag{2.20}$$

Letting $\mathbf{x} = \mathbf{q}$ in (2.17) gives

$$d_t\mathbf{c}(\mathbf{q},t) = -E^\tau\mathbf{c}(\mathbf{q},t) \quad \text{with} \quad \mathbf{c}(\mathbf{q},t)_s = c_s(\mathbf{q},t). \tag{2.21}$$

The matrix E^τ defined by (2.20) may be written in a more useful form by first defining the skew symmetric *coupling matrix* N^τ as

$$N^\tau_{is} = \langle A\mathbf{b}^{\tau,i}, A^{-1}D^j\nabla_{\mathbf{q}}\mathbf{b}^{\tau,s}\mathbf{e}^j - \nabla_{\mathbf{p}}\mathbf{b}^{\tau,s}\nabla_{\mathbf{q}}H_\tau\rangle - \frac{1}{2}\nabla_{\mathbf{q}}\cdot\nabla_{\mathbf{p}}H_\tau\delta_{is}, \tag{2.22}$$

where we observe that this coupling matrix matches the one which appears in [28]. Then (2.20) may be written as

$$E^\tau_{is} = \frac{1}{2}\mathrm{Tr}\left[\nabla_{\mathbf{qp}}H_\tau + M\nabla_{\mathbf{pp}}H_\tau\right]\delta_{is} + N^\tau_{is}. \tag{2.23}$$

The justification for (2.23), which relies on the symmetry of M, is nontrivial and appears in Appendix B.

From (2.23) one may observe a few important properties of the matrix E^τ. First, when the eigenvalue H_τ is simple (not multiple), the skew symmetric N^τ vanishes and E^τ becomes the scalar given by $\frac{1}{2}\mathrm{Tr}\left[\nabla_{\mathbf{qp}}H_\tau + M\nabla_{\mathbf{pp}}H_\tau\right]$. Consequently, from (2.23) we observe that in the case of multiple eigenvalues, all amplitudes $c_s(\mathbf{q},t)$ given in (2.21) evolve as they would in the simple eigenvalue case except for coupling between them determined by N^τ. Furthermore, this coupling is a pure coupling since N^τ is skew symmetric and therefore has purely imaginary eigenvalues. Second, the eigenvalues of E^τ may be explicitly written in terms of the eigenvalues of N^τ. In particular, if the eigenvalues of N^τ are given by λ_i for $i = 1,\dots,r$, then the eigenvalues of E^τ are given by $\frac{1}{2}\mathrm{Tr}\left[\nabla_{\mathbf{qp}}H^\tau + M\nabla_{\mathbf{pp}}H^\tau\right] + \lambda_i$. Finally, this result shows a deep connection between the herein derived Gaussian beam method and the work in [28]. That the matrix N^τ appears in both of these places is perhaps surprising given the very different routes taken to derive it.

In summary, the evolution equations for the Lagrangian Gaussian beams are

$$\begin{aligned}
d_t\mathbf{q} &= \nabla_{\mathbf{p}}H_\tau,\\
d_t\mathbf{p} &= -\nabla_{\mathbf{q}}H_\tau,\\
d_tS &= \nabla_{\mathbf{p}}H_\tau\cdot\mathbf{p} - H_\tau,\\
d_tM &= -\nabla_{\mathbf{qq}}H_\tau - \nabla_{\mathbf{qp}}H_\tau M - M\nabla_{\mathbf{pq}}H_\tau - M\nabla_{\mathbf{pp}}H_\tau M,\\
d_t\mathbf{c} &= -\left\{\tfrac{1}{2}\mathrm{Tr}\left[\nabla_{\mathbf{qp}}H_\tau + M\nabla_{\mathbf{pp}}H_\tau\right]I + N^\tau\right\}\mathbf{c},
\end{aligned} \tag{2.24}$$

with N^τ given by (2.22). As will be explained in section 2.4, the initial conditions for (2.24) are given by

$$\begin{aligned}
\mathbf{q}(\mathbf{q}_0,0) &= \mathbf{q}_0,\\
\mathbf{p}(\mathbf{q}_0,0) &= \nabla_{\mathbf{x}}S_0(\mathbf{q}_0),\\
S(\mathbf{q}_0,0) &= S_0(\mathbf{q}_0),\\
M(\mathbf{q}_0,0) &= \nabla_{\mathbf{xx}}S_0(\mathbf{q}_0) + \tfrac{\mathbf{i}}{\omega}I,\\
c_s(\mathbf{q}_0,0) &= \mathbf{u}_0(\mathbf{q}_0)\cdot[A(\mathbf{q}_0)\mathbf{b}^{\tau,s}(\mathbf{q}_0,\mathbf{p}(\mathbf{q}_0,0))],
\end{aligned} \tag{2.25}$$

where $\omega > 0$ is a real constant.

Remark 2.4. Note that the inclusion of ω in (2.25) differs from the previous Gaussian beam formulation in [14]. Including this parameter is useful in numerical simulations as seen in section 2.3, but it cannot be chosen arbitrarily. Please see the end of section 2.3 for a complete description of the use of ω.

Remark 2.5. For the sake of computing, it may be preferable to write E^τ in the equivalent form

$$E^\tau_{is} = \frac{1}{2}\mathrm{Tr}\left[M\nabla_{\mathbf{pp}}H_\tau\right]\delta_{is} + \langle \mathbf{b}^{\tau,i}, D^j\nabla_{\mathbf{q}}\mathbf{b}^{\tau,s}\mathbf{e}^j - A\nabla_{\mathbf{p}}\mathbf{b}^{\tau,s}\nabla_{\mathbf{q}}H_\tau\rangle \tag{2.26}$$

only because it has fewer terms when expressed this way.

At this stage, we have formulated exactly one Gaussian beam solution to (1.1), but in practice one needs to sum over many Gaussian beam solutions. The details of how to perform this summation are deferred to section 2.4 as we now turn to other matters.

2.2. Conservation of Gaussian profiles. The matrix M, whose governing equation appears in (2.24) and has initial condition given by (2.25), represents the Hessian of the phase of a Gaussian beam, as can be seen from (2.2). In order for the Gaussian beam to have bounded Gaussian profile, M must have positive definite imaginary part. The Gaussian beam can be initialized with positive definite imaginary part (2.25), but further proof is required to show that it remains positive definite for all time. Note that Theorem 2.6 follows directly from the work in [27]. It is reproduced here because some details of the proof are needed elsewhere in this paper.

THEOREM 2.6. *Let $P(t,\mathbf{q}(t,\mathbf{q}_0))$ and $R(t,\mathbf{q}(t,\mathbf{q}_0))$ be the (global) solutions of the equations*

$$\begin{aligned} d_tP &= (\nabla_{\mathbf{pq}}H)P + (\nabla_{\mathbf{pp}}H)R,\\ d_tR &= -(\nabla_{\mathbf{qq}}H)P - (\nabla_{\mathbf{qp}}H)R, \end{aligned} \tag{2.27}$$

with the initial conditions

$$\begin{aligned} P(0,\mathbf{q}_0) &= I,\\ R(0,\mathbf{q}_0) &= M(0,\mathbf{q}_0), \end{aligned} \tag{2.28}$$

where the matrix I is the identity matrix and $\mathrm{Im}(M(0,\mathbf{q}_0))$ is positive definite. Assuming that $M(0,\mathbf{q}_0)$ is symmetric, then for each initial position $\mathbf{q}_0$ the following hold:

1. *$P(t,\mathbf{q}(t,\mathbf{q}_0))$ is invertible for all $t > 0$.*
2. *The solution to the differential equation for M given in (2.24) is given by*

$$M(t,\mathbf{q}(t,\mathbf{q}_0)) = R(t,\mathbf{q}(t,\mathbf{q}_0))P^{-1}(t,\mathbf{q}(t,\mathbf{q}_0)). \tag{2.29}$$

3. *$M(t,\mathbf{q}(t,\mathbf{q}_0))$ is symmetric, and $\mathrm{Im}[M(t,\mathbf{q}(t,\mathbf{q}_0))]$ is positive definite for all $t > 0$.*

Proof. Since $\mathbf{q}(t,\mathbf{q}_0)$ is not directly involved in the proof, simply write $M(t)$, $P(t)$, $R(t)$ to represent the three matrices introduced in the theorem statement.

1. From (2.27), if we let $\mathbf{n} \in \mathbb{C}^n$, then $\mathbf{z}_1 = P(t)\mathbf{n}$ and $\mathbf{z}_2 = R(t)\mathbf{n}$ satisfy

$$\begin{aligned} d_t\mathbf{z}_1 &= (\nabla_{\mathbf{pq}}H)\mathbf{z}_1 + (\nabla_{\mathbf{pp}}H)\mathbf{z}_2,\\ d_t\mathbf{z}_2 &= -(\nabla_{\mathbf{qq}}H)\mathbf{z}_1 - (\nabla_{\mathbf{qp}}H)\mathbf{z}_2. \end{aligned} \tag{2.30}$$

Define

$$\sigma(P,R,\mathbf{n}) = \bar{\mathbf{z}}_1 \cdot \mathbf{z}_2 - \mathbf{z}_1 \cdot \bar{\mathbf{z}}_2. \tag{2.31}$$

Noting that $H(\mathbf{p},\mathbf{q})$ is real, differentiate (2.31) to get

(2.32)

$$\begin{aligned} d_t\sigma(P,R,\mathbf{n}) &= (d_t\bar{\mathbf{z}}_1)\cdot\mathbf{z}_2 + \bar{\mathbf{z}}_1\cdot d_t\mathbf{z}_2 - (d_t\mathbf{z}_1)\cdot\bar{\mathbf{z}}_2 - \mathbf{z}_1\cdot d_t\bar{\mathbf{z}}_2 \\ &= [-(\nabla_{\mathbf{pq}}H)\bar{\mathbf{z}}_1 - (\nabla_{\mathbf{pp}}H)\bar{\mathbf{z}}_2]\cdot\mathbf{z}_2 + \bar{\mathbf{z}}_1\cdot[(\nabla_{\mathbf{qq}}H)\mathbf{z}_1 - (\nabla_{\mathbf{qp}}H)\mathbf{z}_2] \\ &\quad - [(\nabla_{\mathbf{pq}}H)\mathbf{z}_1 + (\nabla_{\mathbf{pp}}H)\mathbf{z}_2]\cdot\bar{\mathbf{z}}_2 - \mathbf{z}_1\cdot[-(\nabla_{\mathbf{qq}}H)\bar{\mathbf{z}}_1 - (\nabla_{\mathbf{qp}}H)\bar{\mathbf{z}}_2] \\ &= 0, \end{aligned}$$

where we use (in the last step) that $\nabla_{\mathbf{pp}}H$ and $\nabla_{\mathbf{qq}}H$ are both symmetric and that $\nabla_{\mathbf{pq}}H = (\nabla_{\mathbf{qp}}H)^T$.

Next assume that $P(t)$ is singular at time $t > 0$. Then let $\mathbf{n} \in \mathbb{C}^n$ be nonzero so that $P(t)\mathbf{n} = 0$. Then one has

$$\begin{aligned} 0 &= \overline{P(t)\mathbf{n}}\cdot R(t)\mathbf{n} - P(t)\mathbf{n}\cdot\overline{R(t)\mathbf{n}} \\ &= \sigma(P(t),R(t),\mathbf{n}) = \sigma(P(0),R(0),\mathbf{n}) \\ &= \overline{P(0)\mathbf{n}}\cdot R(0)\mathbf{n} - P(0)\mathbf{n}\cdot\overline{R(0)\mathbf{n}} \\ &= \bar{\mathbf{n}}\cdot M(0)\mathbf{n} - \mathbf{n}\cdot\overline{M(0)\mathbf{n}} = 2\mathbf{i}\bar{\mathbf{n}}\cdot\mathrm{Im}[M(0)]\mathbf{n}, \end{aligned} \tag{2.33}$$

which contradicts the fact that $\mathrm{Im}[M(0)]$ is positive definite. Thus $P(t)$ is invertible for all $t \geq 0$.

2. Let $M = RP^{-1}$. By differentiating one obtains

$$\begin{aligned} d_tM &= d_t(RP^{-1}) \\ &= (d_tR)P^{-1} + Rd_tP^{-1} \\ &= (d_tR)P^{-1} - RP^{-1}(d_tP)P^{-1} \\ &= [-(\nabla_{\mathbf{qq}}H)P - (\nabla_{\mathbf{qp}}H)R]P^{-1} - RP^{-1}[(\nabla_{\mathbf{pq}}H)P + (\nabla_{\mathbf{pp}}H)R]P^{-1} \\ &= -\nabla_{\mathbf{qq}}H - \nabla_{\mathbf{qp}}HM - M\nabla_{\mathbf{pq}}H - M\nabla_{\mathbf{pp}}HM, \end{aligned} \tag{2.34}$$

which agrees with (2.24).

3. Since both $M(t)$ and its transpose $M(t)^T$ satisfy exactly the same equation, the uniqueness of the solution (see, for example, [5]) implies that $M(t) = M(t)^T$ for all $t > 0$, provided that the initial condition $M(0)$ is symmetric. Next, since $P(t)$ is invertible, take an $\mathbf{n}' \in \mathbb{C}^n$ and define $\mathbf{n} = P(t)^{-1}\mathbf{n}'$ so that

$$\begin{aligned} 2\mathbf{i}\overline{\mathbf{n}'}\cdot\mathrm{Im}[M(t)]\mathbf{n}' &= 2\mathbf{i}\overline{P(t)\mathbf{n}}\cdot\mathrm{Im}[M(t)]P(t)\mathbf{n} \\ &= \overline{P(t)\mathbf{n}}\cdot M(t)P(t)\mathbf{n} - P(t)\mathbf{n}\cdot\overline{M(t)P(t)\mathbf{n}} \\ &= \overline{P(t)\mathbf{n}}\cdot R(t)\mathbf{n} - P(t)\mathbf{n}\cdot\overline{R(t)\mathbf{n}} \\ &= \sigma(P(t),R(t),\mathbf{n}) = \sigma(P(0),R(0),\mathbf{n}) \\ &= \overline{P(0)\mathbf{n}}\cdot R(0)\mathbf{n} - P(0)\mathbf{n}\cdot\overline{R(0)\mathbf{n}} \\ &= \bar{\mathbf{n}}\cdot M(0)\mathbf{n} - \mathbf{n}\cdot\overline{M(0)\mathbf{n}} \\ &= 2\mathbf{i}\bar{\mathbf{n}}\cdot\mathrm{Im}[M(0)]\mathbf{n}. \end{aligned} \tag{2.35}$$

Thus, since $M(0)$ is positive definite, $M(t)$ is also. This completes the proof. ∎

Theorem 2.6 now guarantees that our Gaussian beam solutions of (1.1) remain with a Gaussian profile for all time, provided that they are initialized that way.

2.3. Convergence and Gaussian decomposition. Up until this point, we have constructed just one Gaussian beam solution to (1.1). Here we will prove a few results that show how an arbitrary initial condition of the form shown in (1.1) may be approximated by a sum of many Gaussian beams. Results of this nature have been proved in, for example, [30, 21]. Here, we generalize to the case of symmetric hyperbolic systems. The theoretical results are presented in this section, and the summing process itself is presented in section 2.4.

LEMMA 2.7. *Let $f \in C_0^{j+1}(\mathbb{R}^d \to \mathbb{R})$, and define*

$$v(\mathbf{x},\mathbf{y}) = \left(\frac{1}{2\pi\varepsilon}\right)^{\frac{d}{2}} T_j^y[f](\mathbf{x}) e^{-|\mathbf{x}-\mathbf{y}|^2/2\varepsilon}, \tag{2.36}$$

where $T_j^y[f](\mathbf{x}) = \sum_{|\mathbf{a}|\le j} \frac{(\mathbf{x}-\mathbf{y})^{\mathbf{a}}}{\mathbf{a}!}\partial^{\mathbf{a}} f(\mathbf{y})$ is the jth order Taylor polynomial of $f(\mathbf{x})$ centered at $\mathbf{y}$. Then for $p \ge 1$,

$$\left\| \int_{\mathbb{R}^d} v(\mathbf{x},\mathbf{y}) d\mathbf{y} - f(\mathbf{x}) \right\|_{L^p} \le c\varepsilon^{(j+1)/2}. \tag{2.37}$$

Proof. We introduce the standard multi-index notation wherein if $\mathbf{a} = (a_1, a_2, \ldots, a_d)$ is a d-tuple of nonnegative integers, then define

$$\begin{aligned} |\mathbf{a}| &\equiv a_1 + a_2 + \cdots + a_d, \\ \mathbf{a}! &\equiv a_1! a_2! \cdots a_d!, \\ \partial^{\mathbf{a}} &\equiv \partial_1^{a_1}\partial_2^{a_2}\cdots\partial_d^{a_d}, \\ \mathbf{x}^{\mathbf{a}} &\equiv x_1^{a_1} x_2^{a_2}\cdots x_d^{a_d}. \end{aligned} \tag{2.38}$$

For the above f, Taylor's theorem reads as

$$\begin{aligned} f(\mathbf{x}) &= T_j^y[f](\mathbf{x}) + R_j^y[f](\mathbf{x}) \\ &= \sum_{|\mathbf{a}|\le j} \tfrac{(\mathbf{x}-\mathbf{y})^{\mathbf{a}}}{\mathbf{a}!}\partial^{\mathbf{a}} f(\mathbf{y}) + \sum_{|\mathbf{a}|=j+1} \tfrac{(\mathbf{x}-\mathbf{y})^{\mathbf{a}}}{\mathbf{a}!}\partial^{\mathbf{a}} f(\mathbf{y}+c(\mathbf{x}-\mathbf{y})) \quad \text{for some} \quad c\in(0,1), \end{aligned} \tag{2.39}$$

where we define the remainder term as $R_j^y[f](\mathbf{x}) = \sum_{|\mathbf{a}|=j+1} \frac{(\mathbf{x}-\mathbf{y})^{\mathbf{a}}}{\mathbf{a}!}\partial^{\mathbf{a}} f(\mathbf{y}+c(\mathbf{x}-\mathbf{y}))$. First one can prove a simple upper bound for the remainder:

$$\begin{aligned} |R_j^y[f](x)| &\le \sum_{|\mathbf{a}|=j+1} \left| \tfrac{(\mathbf{x}-\mathbf{y})^{\mathbf{a}}}{\mathbf{a}!}\partial^{\mathbf{a}} f(\mathbf{y}+c(\mathbf{x}-\mathbf{y})) \right| \\ &= |\mathbf{x}-\mathbf{y}|^{|\mathbf{a}|} \sum_{|\mathbf{a}|=j+1} \left| \tfrac{(\mathbf{x}-\mathbf{y})^{\mathbf{a}}}{|\mathbf{x}-\mathbf{y}|^{|\mathbf{a}|}} \tfrac{1}{\mathbf{a}!}\partial^{\mathbf{a}} f(\mathbf{y}+c(\mathbf{x}-\mathbf{y})) \right| \\ &\le |\mathbf{x}-\mathbf{y}|^{j+1} \sum_{|\mathbf{a}|=j+1} \left| \tfrac{1}{\mathbf{a}!}\partial^{\mathbf{a}} f(\mathbf{y}+c(\mathbf{x}-\mathbf{y})) \right| \\ &\le c_0 |\mathbf{x}-\mathbf{y}|^{j+1}, \end{aligned} \tag{2.40}$$

where $c_0 > 0$ is a constant and the last step holds because all of $\partial^{\mathbf{a}} f$ are bounded. Also note that in the above we exclude the point $\mathbf{x} = \mathbf{y}$ where the bound holds trivially. Next we define $g = \left(\frac{1}{2\pi\varepsilon}\right)^{\frac{d}{2}} e^{-|\mathbf{x}-\mathbf{y}|^2/2\varepsilon}$ and then write

$$\begin{aligned} \left\| \int_{\mathbb{R}^d} v(\mathbf{x},\mathbf{y}) d\mathbf{y} - f(\mathbf{x}) \right\|_{L^p} &= \left\| \int_{\mathbb{R}^d} g[T_j^y[f](\mathbf{x}) - f(\mathbf{x})] d\mathbf{y} \right\|_{L^p} \\ &= \left\| \int_{\mathbb{R}^d} g R_j^y[f](\mathbf{x}) d\mathbf{y} \right\|_{L^p}. \end{aligned} \tag{2.41}$$

Since $f(\mathbf{x})$ has compact support, assume that its support is contained in $|\mathbf{x}| < A$, and then note that when both $|\mathbf{x}| > A$ and $|\mathbf{y}| > A$, $R_j^y[f](x) = 0$. Define characteristic functions $\chi_1(\mathbf{x},\mathbf{y}) = \chi_{\{|\mathbf{x}|<A\}\cup\{|\mathbf{y}|<A\}}$, $\chi_2(\mathbf{x}) = \chi_{\{|\mathbf{x}|<2A\}}$, and $\chi_3(\mathbf{x}) = \chi_{\{|\mathbf{x}|\geq 2A\}}$. Then write

(2.42)

$$\begin{aligned} &\left\| \int_{\mathbb{R}^d} g R_j^y[f](\mathbf{x}) d\mathbf{y} \right\|_{L^p} \\ &\quad = \left\| \int_{\mathbb{R}^d} g\chi_1 \left\{ \chi_2 + \tfrac{|\mathbf{x}-\mathbf{y}|^{d+1}}{|\mathbf{x}-\mathbf{y}|^{d+1}} \chi_3 \right\} R_j^y[f](\mathbf{x}) d\mathbf{y} \right\|_{L^p} \\ &\quad \leq \left\| \int_{\mathbb{R}^d} g\chi_1\chi_2 \left| R_j^y[f](\mathbf{x}) \right| d\mathbf{y} + \int_{\mathbb{R}^d} g\chi_1 \tfrac{|\mathbf{x}-\mathbf{y}|^{d+1}}{|\mathbf{x}-\mathbf{y}|^{d+1}} \chi_3 \left| R_j^y[f](\mathbf{x}) \right| d\mathbf{y} \right\|_{L^p} \\ &\quad \leq \left\| \int_{\mathbb{R}^d} g\chi_1\chi_2 c_0 |\mathbf{x}-\mathbf{y}|^{j+1} d\mathbf{y} + \int_{\mathbb{R}^d} g\chi_1 \tfrac{|\mathbf{x}-\mathbf{y}|^{d+1}}{|\mathbf{x}-\mathbf{y}|^{d+1}} \chi_3 c_0 |\mathbf{x}-\mathbf{y}|^{j+1} d\mathbf{y} \right\|_{L^p} \\ &\quad \leq \left\| c_0\chi_2 \int_{\mathbb{R}^d} g|\mathbf{x}-\mathbf{y}|^{j+1} d\mathbf{y} + \int_{\mathbb{R}^d} g\chi_1 \tfrac{|\mathbf{x}-\mathbf{y}|^{d+1}}{||\mathbf{x}|-A|^{d+1}} \chi_3 c_0 |\mathbf{x}-\mathbf{y}|^{j+1} d\mathbf{y} \right\|_{L^p} \\ &\quad \leq \left\| c_0\chi_2 \int_{\mathbb{R}^d} g|\mathbf{x}-\mathbf{y}|^{j+1} d\mathbf{y} + \tfrac{c_0}{||\mathbf{x}|-A|^{d+1}} \int_{\mathbb{R}^d} g\chi_1\chi_3 |\mathbf{x}-\mathbf{y}|^{j+d+2} d\mathbf{y} \right\|_{L^p} \\ &\quad \leq c_0 \left\| \chi_2 \right\|_{L^p} \int_{\mathbb{R}^d} \left(\tfrac{1}{2\pi\varepsilon}\right)^{\frac{d}{2}} e^{-|\mathbf{y}|^2/2\varepsilon} |\mathbf{y}|^{j+1} d\mathbf{y} \\ &\qquad + c_0 \left\| \tfrac{\chi_3}{||\mathbf{x}|-A|^{d+1}} \right\|_{L^p} \int_{\mathbb{R}^d} \left(\tfrac{1}{2\pi\varepsilon}\right)^{\frac{d}{2}} e^{-|\mathbf{y}|^2/2\varepsilon} |\mathbf{y}|^{j+d+2} d\mathbf{y} \\ &\quad \leq c_1 \varepsilon^{\frac{j+1}{2}} + c_2 \varepsilon^{\frac{j+d+2}{2}}, \end{aligned}$$

where c_1 and c_2 are constants. In the fifth line of (2.42), we have used that $|\mathbf{y}| < A$ together with the triangle inequality to get

$$|\mathbf{x}-\mathbf{y}| \geq ||\mathbf{x}| - |\mathbf{y}|| > ||\mathbf{x}| - A|. \tag{2.43}$$

Also, in the seventh line of (2.42) we have used Minkowski's inequality. Finally, for $\varepsilon < 1$ one may redefine the constants to obtain

$$\left\| \int_{\mathbb{R}^d} g R_j^y[f](\mathbf{x}) d\mathbf{y} \right\|_{L^p} \leq c\varepsilon^{\frac{j+1}{2}}. \tag{2.44}$$

This proves the lemma. ∎

THEOREM 2.8. *Let $a \in C_0^{j+1}(\mathbb{R}^d \to \mathbb{R})$ and $\phi \in C_0^{j+3}(\mathbb{R}^d \to \mathbb{R})$. Define*

$$\begin{aligned} u(\mathbf{x}) &= a(\mathbf{x}) e^{\mathbf{i}\phi(\mathbf{x})/\varepsilon}, \\ v(\mathbf{x},\mathbf{y}) &= \left(\tfrac{1}{2\pi\varepsilon\omega}\right)^{\frac{d}{2}} T_j^y[a](\mathbf{x}) e^{\mathbf{i}T_{j+2}^y[\phi](\mathbf{x})/\varepsilon - |\mathbf{x}-\mathbf{y}|^2/2\omega\varepsilon}. \end{aligned} \tag{2.45}$$

Then for $p \geq 1$,

$$\left\| \int_{\mathbb{R}^d} v(\mathbf{x},\mathbf{y}) d\mathbf{y} - u(\mathbf{x}) \right\|_{L^p} \leq c_1 \varepsilon^{-1} (\omega\varepsilon)^{(j+3)/2} + c_2 (\omega\varepsilon)^{(j+1)/2}. \tag{2.46}$$

Proof. As in Lemma 2.7, define $g(\mathbf{x},\mathbf{y}) = \left(\frac{1}{2\pi\omega\varepsilon}\right)^{\frac{d}{2}} e^{-|\mathbf{x}-\mathbf{y}|^2/2\omega\varepsilon}$, and use the notation that $f(\mathbf{x}) = T_j^y[f](\mathbf{x}) + R_j^y[f](\mathbf{x})$, where $R_j^y[f](\mathbf{x})$ is the Lagrangian remainder term from Taylor's theorem given by (2.39). Then one has

$$\begin{aligned}
\left\| \int_{\mathbb{R}^d} v(\mathbf{x},\mathbf{y}) d\mathbf{y} - u(\mathbf{x}) \right\|_{L^p} &= \left\| \int_{\mathbb{R}^d} g \left[T_j^y[a] e^{\mathbf{i} T_{j+2}^y[\phi]/\varepsilon} - a e^{\mathbf{i}\phi/\varepsilon} \right] d\mathbf{y} \right\|_{L^p} \\
&= \left\| \int_{\mathbb{R}^d} g e^{\mathbf{i} T_{j+2}^y[\phi]/\varepsilon} \left[a \left(1 - e^{\mathbf{i} R_{j+2}^y[\phi]/\varepsilon} \right) - R_j^y[a] \right] d\mathbf{y} \right\|_{L^p} \\
&\leq \left\| \int_{\mathbb{R}^d} \left| g \left[a \left(1 - e^{\mathbf{i} R_{j+2}^y[\phi]/\varepsilon} \right) - R_j^y[a] \right] \right| d\mathbf{y} \right\|_{L^p} \\
&\leq \left\| \int_{\mathbb{R}^d} g \left[|a| \left| \left(1 - e^{\mathbf{i} R_{j+2}^y[\phi]/\varepsilon} \right) \right| + \left| R_j^y[a] \right| \right] d\mathbf{y} \right\|_{L^p} \\
&= \left\| \int_{\mathbb{R}^d} g \left[M 2 \left| \sin(R_{j+2}^y[\phi]/2\varepsilon) \right| + \left| R_j^y[a] \right| \right] d\mathbf{y} \right\|_{L^p} \\
&\leq \left\| \int_{\mathbb{R}^d} g \left[M/\varepsilon \left| R_{j+2}^y[\phi] \right| + \left| R_j^y[a] \right| \right] d\mathbf{y} \right\|_{L^p},
\end{aligned} \tag{2.47}$$

where $|a(\mathbf{x})| < M \in \mathbb{R}^+$. Finally, we apply Minkowski's inequality and Lemma 2.7 to obtain that for $\varepsilon < 1$,

$$\begin{aligned}
\left\| \int_{\mathbb{R}^d} v(\mathbf{x},\mathbf{y}) dy - u(\mathbf{x}) \right\|_{L^p} &\leq \left\| \int_{\mathbb{R}^d} g \left[M/\varepsilon \left| R_{j+2}^y[\phi] \right| + \left| R_j^y[a] \right| \right] d\mathbf{y} \right\|_{L^p} \\
&\leq M/\varepsilon \left\| \int_{\mathbb{R}^d} g \left| R_{j+2}^y[\phi] \right| d\mathbf{y} \right\|_{L^p} + \left\| \int_{\mathbb{R}^d} g \left| R_j^y[a] \right| d\mathbf{y} \right\|_{L^p} \\
&\leq c_1 \varepsilon^{-1} (\omega\varepsilon)^{(j+3)/2} + c_2 (\omega\varepsilon)^{(j+1)/2},
\end{aligned} \tag{2.48}$$

which completes the proof. ∎

COROLLARY 2.9. *Let $\mathbf{a}(\mathbf{x}) \in C_0^{j+1}(\mathbb{R}^d \to \mathbb{R}^n)$ and $\phi \in C_0^{j+3}(\mathbb{R}^d \to \mathbb{R})$. Define*

$$\begin{aligned}
\mathbf{u}(\mathbf{x}) &= \mathbf{a}(\mathbf{x}) e^{\mathbf{i}\phi(\mathbf{x})/\varepsilon}, \\
\mathbf{v}(\mathbf{x},\mathbf{y}) &= \left(\tfrac{1}{2\pi\omega\varepsilon}\right)^{\frac{d}{2}} T_j^y[\mathbf{a}](\mathbf{x}) e^{\mathbf{i} T_{j+2}^y[\phi](\mathbf{x})/\varepsilon - |\mathbf{x}-\mathbf{y}|^2/2\omega\varepsilon}.
\end{aligned} \tag{2.49}$$

Then for $p, q \geq 1$,

$$\left\| \left\| \int_{\mathbb{R}^d} \mathbf{v}(\mathbf{x},\mathbf{y}) d\mathbf{y} - \mathbf{u}(\mathbf{x}) \right\|_{l^q} \right\|_{L^p} \leq c_1 \varepsilon^{-1} (\omega\varepsilon)^{(j+3)/2} + c_2 (\omega\varepsilon)^{(j+1)/2}, \tag{2.50}$$

where $\|\mathbf{x}\|_{l^q} \equiv \left(\sum_i |x_i|^q\right)^{1/q}$.

Proof. For any vector $\mathbf{r} \in \mathbb{R}^n$, $\|\mathbf{r}\|_{l^q} \leq \|\mathbf{r}\|_{l^1} = \sum_i |r_i|$. Thus by Minkowski's inequality,

$$\begin{aligned} \left\| \left\| \int_{\mathbb{R}^d} \mathbf{v}(\mathbf{x},\mathbf{y}) d\mathbf{y} - \mathbf{u}(\mathbf{x}) \right\|_{l^q} \right\|_{L^p} &\leq \left\| \sum_i \left| \int_{\mathbb{R}^d} v_i(\mathbf{x},\mathbf{y}) d\mathbf{y} - u_i(\mathbf{x}) \right| \right\|_{L^p} \\ &\leq \sum_i \left\| \int_{\mathbb{R}^d} v_i(\mathbf{x},\mathbf{y}) d\mathbf{y} - u_i(\mathbf{x}) \right\|_{L^p}. \end{aligned} \tag{2.51}$$

The result now follows by repeated use of Theorem 2.8. ∎

These results differ slightly from those in [30, 21] because here we do not use a "cutoff" function and we include the additional constant ω. The parameter ω plays a similar functional role to the cutoff function in that it controls the initial beam width. Even though we have avoided the cutoff function for the above convergence results, one is still free to use it as a postprocessing tool (see [14]). Most importantly, our results are generalized to the symmetric hyperbolic system case.

A function of the ω is to adjust the initial beam width in order to better approximate the initial condition. However, in order for our Gaussian beam method to have the correct asymptotic convergence, ω should be kept as a constant as $\varepsilon \to 0$. The use of ω is not new (see [3, 33, 17]). In particular, [17] discusses how to choose more sophisticated initial Gaussian beam profiles.

In our numerical simulations, we found that $\omega = 1/20$ was sufficient to resolve the initial condition well for the range of ε values used. To clarify our use of ω, we demonstrate numerically the improved approximation of the initial condition with decreasing ω. This demonstration is the first numerical test in section 4. We wish to stress that setting $\omega = 1$ and taking $\varepsilon \to 0$ is perfectly valid, but in some cases, ε is quite small before the initial condition is well resolved enough to obtain a "reasonable" result.

2.4. Lagrangian Gaussian beam summation. Using Corollary 2.9 with $j = 0$ and initial conditions given by (2.25), define

$$\mathbf{u}^{la}_{\varepsilon,\tau}(\mathbf{x},0) \equiv \int_{\mathbb{R}^d} \left(\frac{1}{2\pi\varepsilon\omega} \right)^{d/2} \left(\sum_{s=1}^{r} c_s(\mathbf{q}_0,0) \mathbf{b}^{\tau,s}(\mathbf{q}_0, \mathbf{p}(\mathbf{q}_0,0)) \right) e^{\frac{\mathbf{i}}{\varepsilon} T(\mathbf{x},\mathbf{q}_0,0)} d\mathbf{q}_0, \tag{2.52}$$

where T is given by (2.2). Using the evolution equations for $\mathbf{q}, \mathbf{p}, S, M$, and $\mathbf{c}$ given by (2.24) along with initial conditions given by (2.25), the discrete version of (2.52) at a time $t > 0$ then reads as

$$\mathbf{u}^{la}_{\varepsilon,\tau}(\mathbf{x},t) = \sum_{\mathbf{q}^i_0} \left(\frac{1}{2\pi\varepsilon\omega} \right)^{d/2} \left(\sum_{s=1}^{r} c_s(\mathbf{q}^i,t) \mathbf{b}^{\tau,s}(\mathbf{q}^i, \mathbf{p}(\mathbf{q}^i,t)) \right) e^{\frac{\mathbf{i}}{\varepsilon} T(\mathbf{x},\mathbf{q}^i,t)} \left| \Delta \mathbf{q}^i_0 \right|, \tag{2.53}$$

where $\mathbf{q}^i \equiv \mathbf{q}(\mathbf{q}^i_0, t)$, and $\mathbf{q}^i_0$ is an equidistant grid of points in the domain with grid spacing $\left|\Delta \mathbf{q}^i_0\right|$. Equation (2.53) is the summation formula for our Lagrangian Gaussian beam method corresponding to the eigenvalue H_τ. It may be performed at any time $t > 0$. Performing the derivation in section 2.1 for each H_τ and summing the results gives

$$\mathbf{u}^{la}_{\varepsilon} = \sum_{\tau} \mathbf{u}^{la}_{\varepsilon,\tau}, \tag{2.54}$$

which is our Lagrangian Gaussian beam approximate solution to (1.1). Finally, from section 2.3, we have

$$\left\| \left\| \mathbf{u}_\varepsilon^{la}(\mathbf{x},0) - \mathbf{u}_\varepsilon(\mathbf{x},0) \right\|_{l^q} \right\|_{L^p} \leq c_1 \varepsilon^{-1} (\omega\varepsilon)^{3/2} + c_2 (\omega\varepsilon)^{1/2}. \tag{2.55}$$

This completes the Lagrangian formulation of Gaussian beams, and we proceed to the Eulerian formulation in the next section.

3. The Eulerian formulation. In the Lagrangian formulation of Gaussian beams, individual beams, which may be evenly spaced over the domain when initialized, tend to spread apart over time. If one wishes to avoid this phenomenon, they may employ an Eulerian formulation of Gaussian beams. We present one such Eulerian formulation here.

Start by forming a vector function $\mathbf{\Phi}(t,\mathbf{p},\mathbf{q})$ on phase space whose real part tracks the evolution of each component of $\mathbf{p}$ and $\mathbf{q}$ by taking

$$\mathcal{L}\mathbf{\Phi} = 0 \quad \text{with} \quad \mathbf{\Phi}(0,\mathbf{p},\mathbf{q}) = [\mathbf{p} - \nabla_{\mathbf{q}} S_0(\mathbf{q})] - \tfrac{\mathbf{i}}{\omega}\mathbf{q}, \tag{3.1}$$

where

$$\mathcal{L} \equiv \partial_t + \nabla_{\mathbf{p}} H_\tau \cdot \nabla_{\mathbf{q}} - \nabla_{\mathbf{q}} H_\tau \cdot \nabla_{\mathbf{p}}. \tag{3.2}$$

With the above vector function, a quick derivation gives the evolution equations and initial conditions for the matrices $-\nabla_{\mathbf{q}}\mathbf{\Phi}$ and $\nabla_{\mathbf{p}}\mathbf{\Phi}$ as

$$\begin{aligned}
\mathcal{L}(-\nabla_{\mathbf{q}}\mathbf{\Phi}) &= -(\nabla_{\mathbf{pq}} H_\tau)(-\nabla_{\mathbf{q}}\mathbf{\Phi}) - (\nabla_{\mathbf{qq}} H_\tau)(\nabla_{\mathbf{p}}\mathbf{\Phi}) \\
&\qquad \text{with} \quad -\nabla_{\mathbf{q}}\mathbf{\Phi}(0,\mathbf{p},\mathbf{q}) = \nabla_{\mathbf{qq}} S_0(\mathbf{q}) + \tfrac{\mathbf{i}}{\omega} I, \\
\mathcal{L}(\nabla_{\mathbf{p}}\mathbf{\Phi}) &= (\nabla_{\mathbf{pp}} H_\tau)(-\nabla_{\mathbf{q}}\mathbf{\Phi}) + (\nabla_{\mathbf{qp}} H_\tau)(\nabla_{\mathbf{p}}\mathbf{\Phi}) \\
&\qquad \text{with} \quad \nabla_{\mathbf{p}}\mathbf{\Phi}(0,\mathbf{p},\mathbf{q}) = I,
\end{aligned} \tag{3.3}$$

which are, by construction, the phase space equivalents to (2.27). As noticed in [14], one may compute $M(t,\mathbf{q},\mathbf{p})$ using the formula

$$M = (-\nabla_{\mathbf{q}}\mathbf{\Phi})(\nabla_{\mathbf{p}}\mathbf{\Phi})^{-1}, \tag{3.4}$$

where the initial condition for M given in (2.25) is satisfied by our choice of initial condition for $\mathbf{\Phi}$ shown in (3.1). It is important to note that one could just as well solve the phase space equation for M which follows directly from (2.24), but by using (3.1) instead, we trade solving a matrix of coupled Liouville equations for solving a vector of homogeneous (uncoupled) Liouville equations.

Finally, the phase space equations for S and $\mathbf{c}$ follow directly from (2.24) so that, in summary, we obtain the following collection of Liouville equations which defines the Eulerian Gaussian beam formulation:

$$\begin{aligned}
\mathcal{L}\mathbf{\Phi} &= 0 \\
&\qquad \text{with} \quad \mathbf{\Phi}(0,\mathbf{q},\mathbf{p}) = [\mathbf{p} - \nabla_{\mathbf{q}} S_0(\mathbf{q})] - \tfrac{\mathbf{i}}{\omega}\mathbf{q}, \\
\mathcal{L}S &= \nabla_{\mathbf{p}} H_\tau \cdot \mathbf{p} - H_\tau \\
&\qquad \text{with} \quad S(0,\mathbf{q},\mathbf{p}) = S_0(\mathbf{q}), \\
\mathcal{L}\mathbf{c} &= -\left\{ \tfrac{1}{2} \mathrm{Tr}\left[\nabla_{\mathbf{qp}} H_\tau + M \nabla_{\mathbf{pp}} H_\tau \right] I + N^\tau \right\} \mathbf{c} \\
&\qquad \text{with} \quad c_s(0,\mathbf{q},\mathbf{p}) = \mathbf{u}_0(\mathbf{q}) \cdot [A(\mathbf{q})\mathbf{b}^{\tau,s}(\mathbf{q},\mathbf{p})],
\end{aligned} \tag{3.5}$$

where N is the matrix given by (2.20) with M replaced by $(-\nabla_{\mathbf{q}}\mathbf{\Phi})(\nabla_{\mathbf{p}}\mathbf{\Phi})^{-1}$ wherever it appears in accordance with (3.4).

The Eulerian summation formula for eigenvalue H_τ follows from (2.52) in the same way as it did in [14]. It reads as

(3.6)
$$\mathbf{u}^{eu}_{\varepsilon,\tau}(\mathbf{x},t) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left(\frac{1}{2\pi\varepsilon\omega}\right)^{\frac{d}{2}}\sum_{s=1}^{r} c_s(t,\mathbf{q},\mathbf{p})\mathbf{b}^{\tau,s}(\mathbf{q},\mathbf{p})e^{\frac{\mathbf{i}}{\varepsilon}T(t,\mathbf{x},\mathbf{q},\mathbf{p})}\delta(\mathrm{Re}[\mathbf{\Phi}(t,\mathbf{q},\mathbf{p})])d\mathbf{p}d\mathbf{q}.$$

Note that using (3.6) correctly is a subtle matter which we will discuss in section 3.2. Performing the above derivation for each H_τ and summing the results gives

$$\mathbf{u}^{eu}_{\varepsilon} = \sum_{\tau}\mathbf{u}^{eu}_{\varepsilon,\tau}. \tag{3.7}$$

Equation (3.7) is our Eulerian Gaussian beam approximate solution to (1.1).

3.1. One-dimensional simplifications for Eulerian Gaussian beams. In general, the disadvantage of the Eulerian formulation as compared to the Lagrangian formulation is that it requires solving the Liouville equations (3.5) on $2d$-dimensional phase space. In general, one may take advantage of optimized numerical solvers to help mitigate this computational cost (see, for example, [24]). However, when $d = 1$, we may take advantage of the following simplifications.

For one-dimensional systems of the form (1.1), it is straightforward to show that the eigenvalues of the dispersion matrix (2.7) will always be of the form $H_\tau(q,p) = pf(q)$ for some function $f(q)$. As a consequence, the eigenvectors are independent of p. With the help of the following simple theorem, we can make use of these facts to reduce the system (3.5) to one-dimensional computations. This trick was first used in [11].

THEOREM 3.1. *The solution to*

$$\begin{cases}[\partial_t + f(q)\partial_q - pf'(q)\partial_p]g = 0,\\ g(q,p,0) = p - \partial_q S_0(q)\end{cases} \tag{3.8}$$

may be written as $g(t,q,p) = p\Gamma_1(t,q) + \Gamma_0(t,q)$ *with* Γ_0 *and* Γ_1 *governed by*

$$\begin{cases}[\partial_t + f(q)\partial_q]\Gamma_0 = 0,\\ \Gamma_0(q,0) = -\partial_q S_0(q)\end{cases} \tag{3.9}$$

and

$$\begin{cases}[\partial_t + f(q)\partial_q - f'(q)]\Gamma_1 = 0,\\ \Gamma_1(q,0) = 1.\end{cases} \tag{3.10}$$

Proof. The proof is a simple substitution of the assumed form for g into (3.8). ∎

With the above theorem, we may write the evolution equations for Eulerian Gaussian beams as

$$\begin{aligned}
&[\partial_t + f(q)\partial_q]\Gamma_0 = 0, && \text{with } \Gamma_0(0,q) = -\partial_q S_0(q) - \tfrac{\mathbf{i}}{\omega}q,\\
&[\partial_t + f(q)\partial_q]\Gamma_1 = f'(q)\Gamma_1, && \text{with } \Gamma_1(0,q) = 1,\\
&[\partial_t + f(q)\partial_q]S = 0, && \text{with } S(0,q) = S_0(q),\\
&[\partial_t + f(q)\partial_q]\mathbf{c} = -E^\tau\mathbf{c}, && \text{with } c_i(0,q,p) = \mathbf{u}_0(q)\cdot[A(q)\mathbf{b}^{\tau,i}(q)],
\end{aligned} \tag{3.11}$$

where $\Phi(t,q,p) = p\Gamma_1 + \Gamma_0$ and where E^τ simplifies to

$$E^\tau_{is} = \langle \mathbf{b}^{\tau,i}(q), D\partial_q \mathbf{b}^{\tau,s}(q)\rangle. \tag{3.12}$$

Solving (3.11) is now a one-dimensional computation which greatly increases efficiency. The existence of an extension of the decomposition in Theorem 3.1 to higher-dimensional space remains an open question.

3.2. New Eulerian summation formula. An Eulerian formulation for Gaussian beams is not a new idea, but in such papers as [14], the Eulerian formulation lost accuracy after the formation of caustics, and a semi-Lagrangian computation was used there to avoid this problem. Here we introduce a new fully Eulerian summation formula that also avoids this problem.

Observe that we may remove the delta-function from (3.6) by integrating over any of the d coordinates out of the total $2d$ coordinates on phase space (since $\mathrm{Re}[\mathbf{\Phi}(t,\mathbf{q},\mathbf{p})]$ is a d-dimensional vector), and the proper choice is the one which is least singular. To illustrate this point, examine the one-dimensional case which has two-dimensional phase space wherein one obtains

$$\mathbf{u}^{eu}_{\varepsilon,\tau}(x,t) = \sum_{s=1}^{r} \int_{\mathbb{R}}\int_{\mathbb{R}} \left(\frac{1}{2\pi\varepsilon\omega}\right)^{\frac{d}{2}} c(t,q,p)\mathbf{b}^{\tau,s}(q,p) e^{\frac{\mathbf{i}}{\varepsilon}T(t,x,q,p)} \delta(\mathrm{Re}[\Phi(t,q,p)])dqdp. \tag{3.13}$$

This may be written as

$$\mathbf{u}^{eu}_{\varepsilon,\tau}(x,t) = \sum_{s=1}^{r}\sum_{p_k(q)} \int_{\mathbb{R}} \left(\frac{1}{2\pi\varepsilon\omega}\right)^{\frac{d}{2}} c(t,q,p_k(q))\mathbf{b}^{\tau,s}(q,p_k(q)) \frac{e^{\frac{\mathbf{i}}{\varepsilon}T(t,x,q,p_k(q))}}{|\mathrm{Re}[\partial_p\Phi(t,q,p_k(q))]|}dq \tag{3.14}$$

or

$$\mathbf{u}^{eu}_{\varepsilon,\tau}(x,t) = \sum_{s=1}^{r}\sum_{q_k(p)} \int_{\mathbb{R}} \left(\frac{1}{2\pi\varepsilon\omega}\right)^{\frac{d}{2}} c(t,q_k(p),p)\mathbf{b}^{\tau,s}(q_k(p),p) \frac{e^{\frac{\mathbf{i}}{\varepsilon}T(t,x,q_k(p),p)}}{|\mathrm{Re}[\partial_q\Phi(t,q_k(p),p)]|}dp, \tag{3.15}$$

where $p_k(q)$ and $q_k(p)$ enumerate (in k) the points where $\mathrm{Re}\,[\mathbf{\Phi}(t,q,p)]$ vanishes for a given q or p value, respectively. Note that the value of k in $p_k(q)$ and $q_k(p)$ can change for a given choice of q or p, respectively. In particular, k is often larger than 1 when a solution becomes multivalued. The points $p_k(q)$ and $q_k(p)$ may be found by interpolation of the surface $\mathrm{Re}\,[\mathbf{\Phi}(t,q,p)]$ (see Remark 3.3). Either summation formula, (3.14) or (3.15), is valid, but it makes sense to locally choose the less singular of the two. To this end define

$$\Gamma_q(a,b) = |\mathrm{Re}[\partial_q\mathbf{\Phi}(t,q,p)]| \quad \text{and} \quad \Gamma_p(a,b) = |\mathrm{Re}[\partial_p\mathbf{\Phi}(t,q,p)]|, \tag{3.16}$$

and let (q_i,p_j) be a rectangular grid on phase space with spacing h_q and h_p in the q

and p coordinates, respectively. Then take

$$\begin{aligned}\mathbf{u}^{eu}_{\varepsilon,\tau}(x,t) = h_p \sum_{s=1}^{r} \sum_{q_i} \sum_{\{p_k(q_i)|\Gamma_p(q_i,p_k(q_i))>\Gamma_q(q_i,p_k(q_i))\}} \\ \left[\left(\tfrac{1}{2\pi\varepsilon\omega}\right)^{\frac{d}{2}} c(t,q_i,p_k(q_i))\mathbf{b}^{\tau,s}(q_i,p_k(q_i)) \tfrac{e^{\frac{\mathbf{i}}{\varepsilon}T(t,x,q_i,p_k(q_i))}}{\Gamma_p(q_i,p_k(q_i))} \right] \\ + h_q \sum_{s=1}^{r} \sum_{p_i} \sum_{\{q_k(p_i)|\Gamma_p(q_k(p_i),p_i)<\Gamma_q(q_k(q_i),p_i)\}} \\ \left[\left(\tfrac{1}{2\pi\varepsilon\omega}\right)^{\frac{d}{2}} c(t,q_k(p_i),p_i)\mathbf{b}^{\tau,s}(q_k(p_i),p_i) \tfrac{e^{\frac{\mathbf{i}}{\varepsilon}T(t,x,q_k(p_i),p_i)}}{\Gamma_q(q_k(p_i),p_i)} \right].\end{aligned} \tag{3.17}$$

This approach naturally avoids the lack of resolution near caustics, is fully Eulerian, and, though appearing complicated when written out, is actually quite intuitive to implement in code. It may also be extended to higher dimensions (see section 3.4) and requires no manual tracking of the caustics. Notice that one could also change coordinates in (3.13) and remove the delta-function using arbitrary coordinates, but since we will be solving our Liouville equations on rectangularly gridded phase space, our proposed approach is the easiest. As far as we are aware, a fully Eulerian method such as this has not been introduced before. In the next section we recreate the experiment in [14] to demonstrate the usefulness of this new approach.

3.3. New summation formula for the Schrödinger equation. In this section, we present an illustrative example to demonstrate the effectiveness of our new summation formula. One can show that Theorem 2.8 together with the initial conditions for Γ_1 specified by (3.11) implies that $\Gamma_1 > 0$ for all $t \geq 0$, which implies that caustics will not form in the one-dimensional hyperbolic system. Because it is precisely when caustics form that our new summation formula (3.17) is most effective, we would need at least a two-dimensional hyperbolic system with four-dimensional phase space, which is hard to visualize and serves as a poor illustration of our method. Thus, instead we present the following example involving the Schrödinger equation which demonstrates not only how the method works but also that it may be applied to Eulerian Gaussian beams for a wide class of problems beyond the symmetric hyperbolic systems studied in this paper.

The example we take is Example 3 from [14]. Referring to [14] for the background, we take all parameters to be identical except instead of using $\varepsilon = 1/10000$, we will use $\varepsilon = 1/5000$ because the images are clearer. Just as in [14], we use the time splitting spectral method for the reference solution. After adjusting our new summation formula (3.17) for the Schrödinger equation case, we compare the old and new summation formulas in Figure 1. The new method clearly fares much better.

Remark 3.2. Each solution shown in Figure 1 was computed with exactly the same data on the phase space with the difference being entirely in the postprocessing. Also [14] did present a solution for the errors seen in Figure 1. However, the idea there was to find the caustics manually, add beams near the caustics, and solve for them using a semi-Lagrangian technique. Thus the method there was not truly Eulerian, whereas our proposed method is.

To better understand the difference between (3.14) and (3.17), we note that the two parts of (3.17) correspond to points in phase space which are evenly distributed along the q-axis ("vertically" summed) and points evenly spaced along the p-axis ("horizontally" summed). The difference becomes evident when we look at the actual points along the zero set of the real part of $\mathbf{\Phi}$ that each method is using. This is

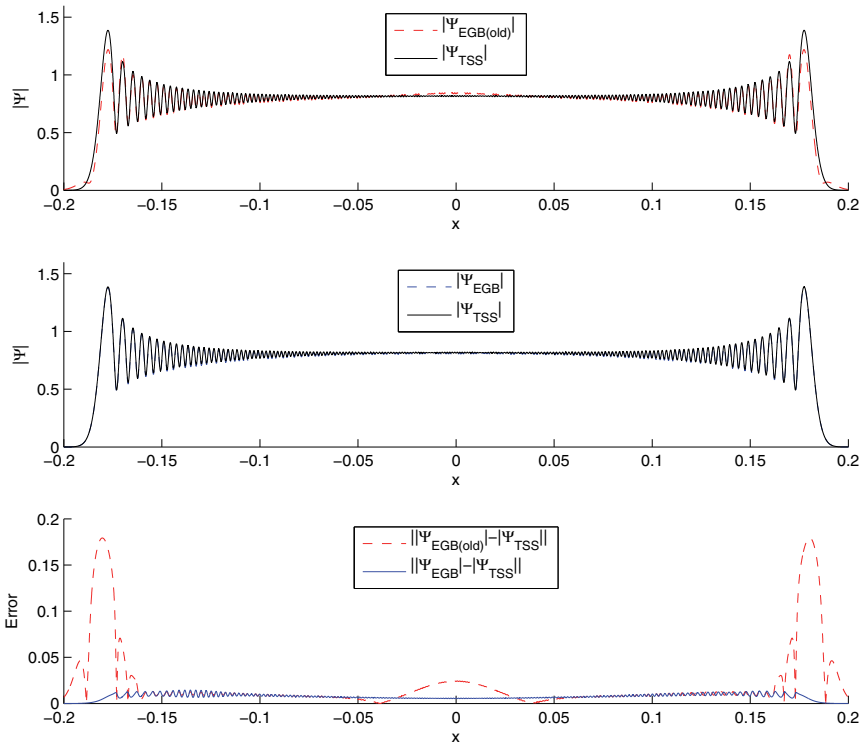

FIG. 1. *Top: Comparison of the Eulerian Gaussian beam (EGB) solution using the old summing method* (3.14) *with the reference solution computed with the time splitting spectral method (TSS). Middle: Comparison of the solution using the new summing method* (3.17) *with the reference solution. Bottom: Comparison of the errors of each method.*

depicted in Figure 2. Observe that the zero contour is well resolved even at the caustics which are near $q = \pm.18$.

Remark 3.3. In order to apply (3.17) one must numerically detect the zero set of $\mathrm{Re}\,[\mathbf{\Phi}(t,q,p)]$. In order to get any kind of reasonable results for the simulations which resulted in Figure 1, one must use a method to find the zero set which is at least second order in the step size. Our implementation used linear interpolation of the surface $\mathrm{Re}\,[\mathbf{\Phi}(t,q,p)]$ to approximate points on the zero set.

3.4. New summation formula in higher dimension: A sketch. The idea used in the new summation formula (3.17) may be extended to higher dimensions. Although we have not implemented it in this paper, we present the idea.

In $2d$-dimensional phase space we have the $2d$ coordinates $(\mathbf{q},\mathbf{p})$, and we denote the subsets of d coordinates as $S^j = (s_1^j,\ldots,s_n^j)$ where each s_i^j equals some coordinate in $(\mathbf{q},\mathbf{p})$. Noting that there are $\binom{2d}{d}$ of these subsets, we then compute the following volumes at each point on the zero set:

$$V^j = \sqrt{\det(\nabla_{S^j}\mathrm{Re}(\mathbf{\Phi}))}, \tag{3.18}$$

where ∇_{S^j} is the gradient with respect to the coordinates $(s_1^j,\ldots,s_n^j)$. For points where $V^m \le V^j$ for all $j = 1,\ldots,d$, we perform the Eulerian sum in the coordinates

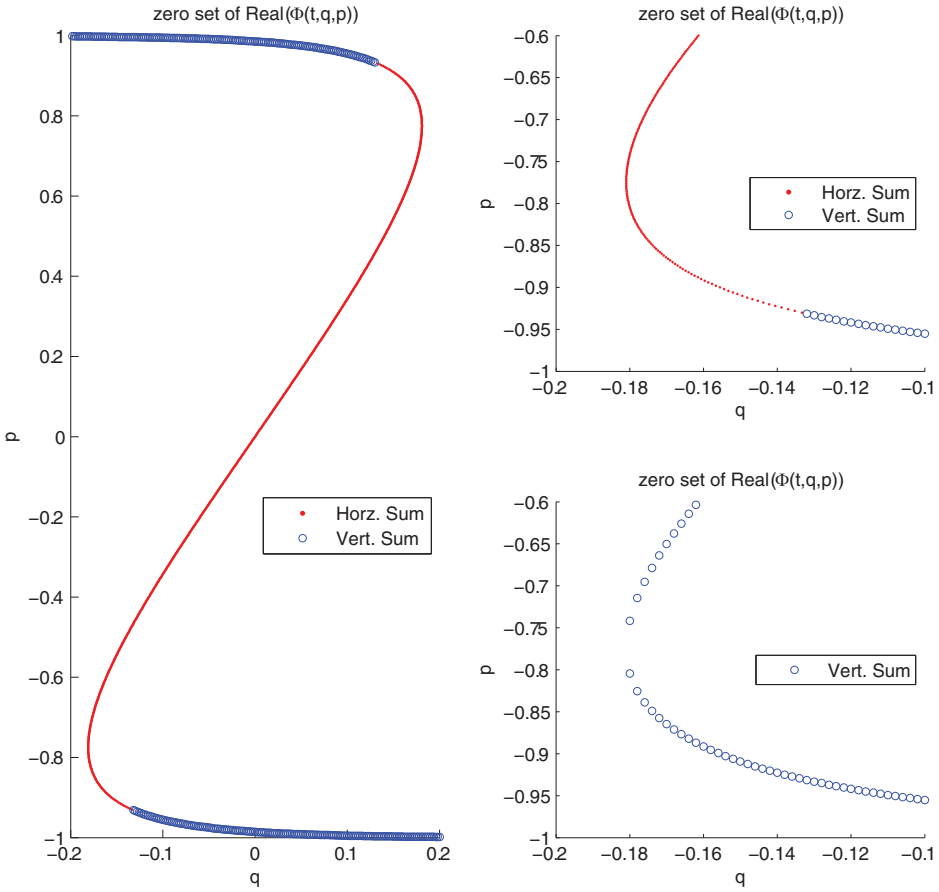

FIG. 2. *Left: The zero set along with the points selected by the new summation method* (3.17) *separated by "vertical" summing and "horizontal" summing. Right top: This is a zoomed-in plot of the left-hand plot. Right bottom: This is the same zoomed-in plot as the right top plot but using the points selected from the old summation method* (3.14)*. In all images, the caustics appear near* $q = \pm.18$*, and, as is seen, with the old method (right bottom plot) the caustics are not well resolved by the selected points.*

S^m. This way, the zero set of $\mathrm{Re}(\mathbf{\Phi})$ is naturally divided into regions where each set of coordinates S^j is used for the summation. This approach avoids any singularities in the summation.

Finding the zero set in higher dimension is complicated. To see how it might be done, consider the $d = 2$ case. We want to find the intersection of the zero sets of $\mathrm{Re}\,[\phi_1(t, \mathbf{q}, \mathbf{p})]$ and $\mathrm{Re}\,[\phi_2(t, \mathbf{q}, \mathbf{p})]$. Without loss of generality, assume our subset of coordinates is $\mathbf{q}$, so our goal is, given a $\mathbf{q}$, to find all $\mathbf{p}$ where both $\mathrm{Re}\,[\phi_1(t, \mathbf{q}, \mathbf{p})]$ and $\mathrm{Re}\,[\phi_2(t, \mathbf{q}, \mathbf{p})]$ are zero. Given $\mathbf{q}$, we fix p_1 and then vary p_2 to find all values p_2 where $\mathrm{Re}\,[\phi_1(t, \mathbf{q}, \mathbf{p})] = 0$. Finding these points is just a matter of looking for where the sign of $\mathrm{Re}\,[\phi_1(t, \mathbf{q}, \mathbf{p})]$ changes as we vary p_2 over the Eulerian grid. Thus we can formulate a possibly multivalued function $p_2(p_1)$ which returns all values of p_2 found in (for example) increasing order. Then for each branch of this multivalued function, we recursively use the same process on the function $\mathrm{Re}\,[\phi_2(t, \mathbf{q}, (p_1, p_2(p_1)))]$ which

is now a one-dimensional function in p_1. We can now vary p_1 to find the zeros of $\mathrm{Re}\,[\phi_2(t, \mathbf{q}, (p_1, p_2(p_1)))]$ which will give all the points on the intersection of the zero sets of $\mathrm{Re}\,[\phi_1(t, \mathbf{q}, \mathbf{p})]$ and $\mathrm{Re}\,[\phi_2(t, \mathbf{q}, \mathbf{p})]$ for each $\mathbf{q}$. For $d > 2$ dimensions, we can use the same idea, where each recursive step drops the dimension of the search by one until points are identified. An algorithm to perform the above may be implemented in any programming language using recursive function calls. However, we have not implemented the higher-dimensional case in this paper. Doing this efficiently may be the subject of future work.

4. Numerical results. In this section we present our numerical results. These examples include solutions to one- and two-dimensional systems, both Lagrangian and Eulerian formulations as well as convergence results. As was discussed at the end of section 2.3, the parameter ω should be chosen so that the initial Gaussian beam decomposition is sufficiently accurate. This step was performed for all the following numerical examples before the simulations were performed with various decreasing values ε. The values chosen for ω are indicated in the parameters for each experiment. A demonstration of this process is presented in the first numerical simulation of section 4.1.

4.1. One-dimensional system. This numerical experiment will test the non-strict hyperbolic case in one dimension. We use the following one-dimensional system which we also used in [11].

In reference to (1.1) define $D^1 = I$ and $A(x)$ by

$$A^{-1} = RKR^T \tag{4.1}$$

with

$$K = \begin{pmatrix} a & b & b \\ b & a & b \\ b & b & a \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4.2}$$

where $0 < b < a$ and a, b, θ are functions of x. The eigenvalues and eigenvectors of the dispersion matrix L which are orthonormal with respect to $\langle \cdot, \cdot \rangle_A$ (defined by (2.9)) are

(4.3)
$$\begin{cases} H_1 = k(a-b), & \mathbf{b}^{1,1} = \sqrt{a-b}R\mathbf{v}_{1,1}, \\ & \mathbf{b}^{1,2} = \sqrt{a-b}R\mathbf{v}_{1,2}, \\ H_2 = k(a+2b), & \mathbf{b}^{2} = \sqrt{a+2b}R\mathbf{v}_{2}, \end{cases} \quad \text{where} \quad \begin{cases} \mathbf{v}^{1,1} = \frac{1}{\sqrt{2}}(1,0,-1), \\ \mathbf{v}^{1,2} = \frac{1}{\sqrt{6}}(1,-2,1), \\ \mathbf{v}^{2} = \frac{1}{\sqrt{3}}(1,1,1). \end{cases}$$

The coupling matrix (2.22) for H_1 is

$$N^1 = \begin{pmatrix} 0 & \frac{(a-b)\theta'}{\sqrt{3}} \\ -\frac{(a-b)\theta'}{\sqrt{3}} & 0 \end{pmatrix}. \tag{4.4}$$

First we demonstrate our use of the ω parameter to improve the Gaussian beam decomposition of the initial condition. Take the following parameter.

Parameters 4.1. Take $a(x) = 2$, $b(x) = 1$, and $\theta(x) = 10\sin(2\pi x)$. The domain is $x \in [-.5, .5]$. The initial conditions are $\mathbf{u}_0(x) = (e^{-40\tan(\pi x)^2}, 0, 0)$, $S_0(x) = x$. Take

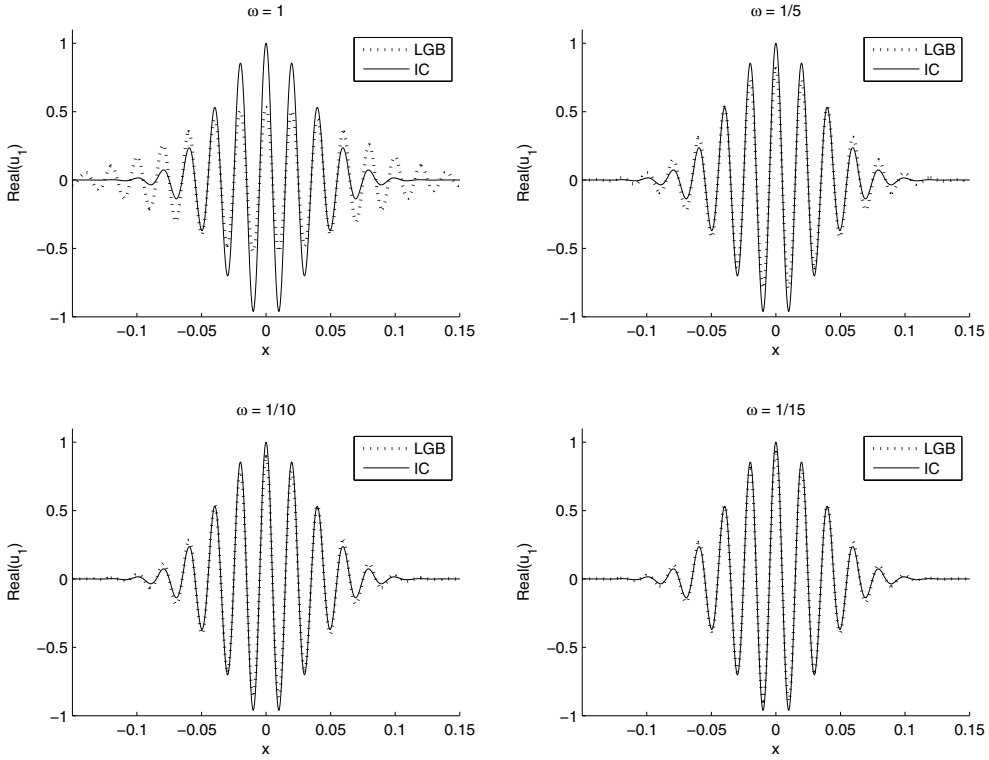

FIG. 3. *Comparison of Lagrangian Gaussian beam (LGB) decomposition to the initial condition (IC) for Parameters* 4.1. *Plotted are the real parts of the first component of the solution.*

$\varepsilon = 1/(100\pi)$, $\omega = 1, 1/5, 1/10, 1/15$, and final time $t_f = 0$. The number of beams is 150 evenly spaced over the domain.

Because $t_f = 0$, in this simulation we decompose the initial condition into Gaussian beams, and then immediately sum them up again. Figure 3 compares the result to the initial condition for various values of ω. The decomposition of the initial condition improves with decreasing ω as expected.

Next we test the full method with $t_f > 0$ by performing a simulation with the following parameters.

Parameters 4.2. Take $a(x) = 2$, $b(x) = 1$, and $\theta(x) = 10\sin(2\pi x)$. The domain is $x \in [-.5, .5]$. The initial conditions are $\mathbf{u}_0(x) = (e^{-40\tan(\pi x)^2}, 0, 0)$, $S_0(x) = x$. Take $\varepsilon = 1/(500\pi)$, $\omega = 1/20$, and final time $t_f = .05$. The number of beams is 150 evenly spaced over the domain.

The result is illustrated in Figure 4 and shows good agreement.

For the above case when both a and b are constant in the matrix (4.2), many terms of the Gaussian beam ODE system given by (2.24) vanish. For a numerical test where these terms don't vanish, take the following.

Parameters 4.3. Take $a = 2(.5 + .4\sin(4\pi x))$, $b = (.5 + .4\sin(4\pi x))$, and $\theta(x) = \sin(2\pi x)$. The domain is $x \in [-.5, .5]$. The initial conditions are $\mathbf{u}_0(x) = (e^{-40\tan(\pi x)^2}, 0, 0)$, $S_0(x) = x$. Let $\varepsilon = 1/(500\pi)$ and $\omega = 1/20$. The final time is $t_f = .05$. The number of beams is 150 evenly spaced over the domain.

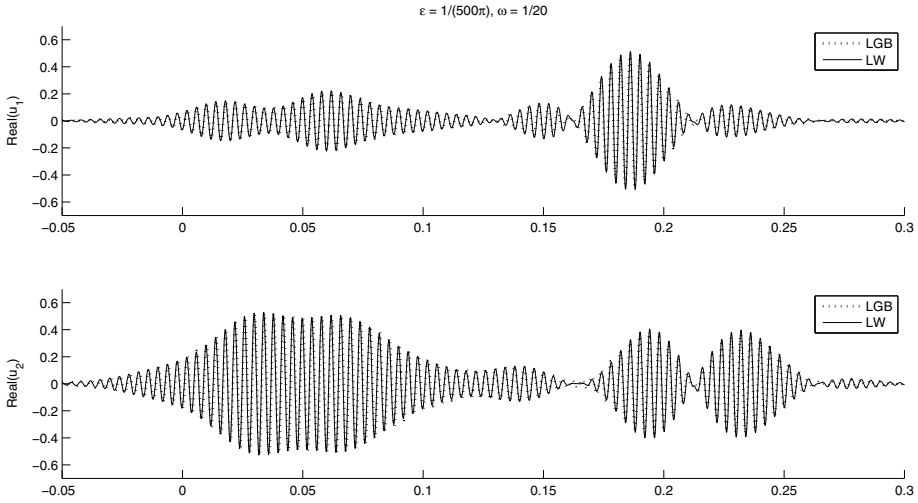

FIG. 4. *Comparison of Lagrangian Gaussian beam method with a reference solution for Parameters* 4.2. *Plotted are the real parts of the first two components of the one-dimensional system with constant* a *and* b.

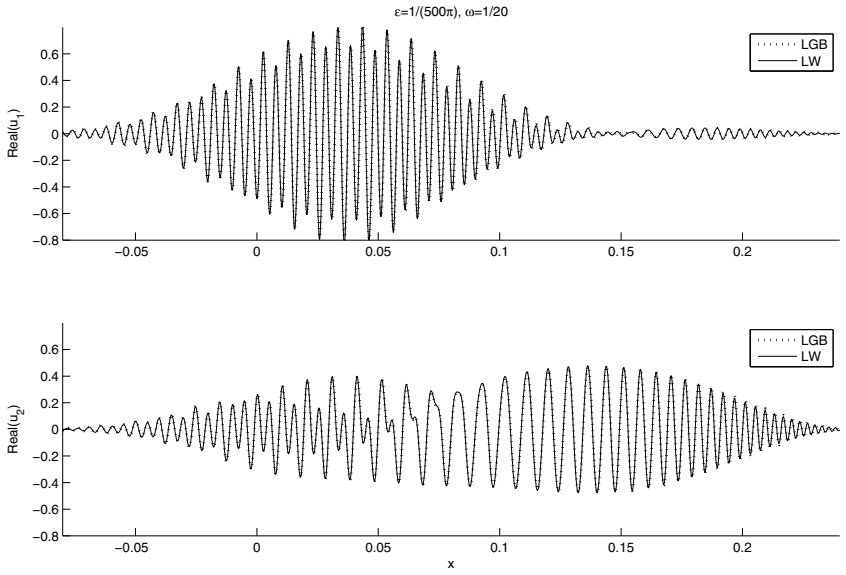

FIG. 5. *Comparison of Lagrangian Gaussian beam method with a reference solution for Parameters* 4.3. *Plotted are the real parts of the first two components of the one-dimensional system with nonconstant* $a(x)$ *and* $b(x)$.

The result is shown in Figure 5, and the agreement is again good.

We now verify numerically the Eulerian formulation using our new summation method discussed in section 3.2. For this example, we use the same one-dimensional system (4.2) with the following parameters.

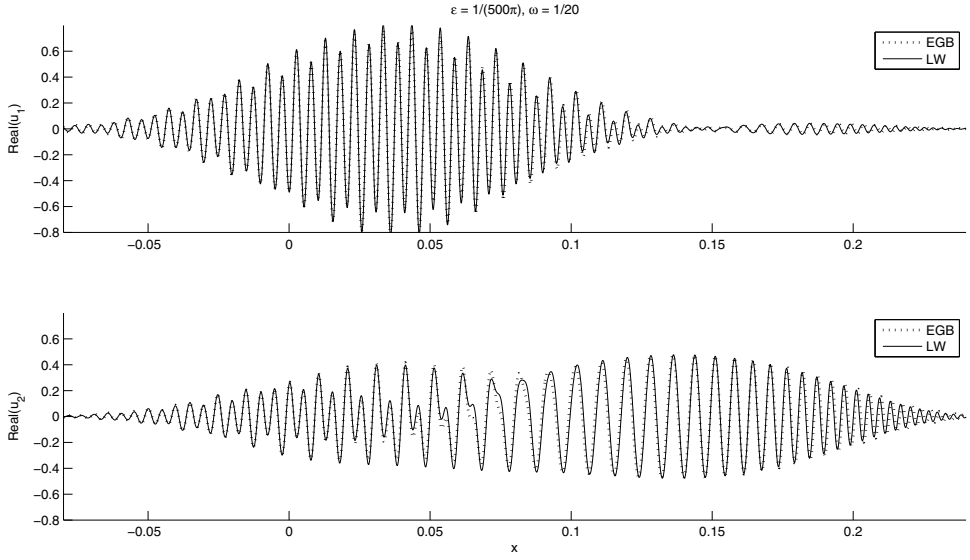

FIG. 6. *Comparison of Eulerian Gaussian beam with the exact solution for Parameters* 4.4. *Plotted are the real parts of the first two components of the one-dimensional system with nonconstant* $a(x)$ *and* $b(x)$.

TABLE 1
Convergence data for Lagrangian Gaussian beam method.

ε	$\frac{1}{32\pi}$	$\frac{1}{64\pi}$	$\frac{1}{128\pi}$	$\frac{1}{256\pi}$	
Beam count	60	80	110	160	Convergence rate
L^1 error	0.0286396	0.0186261	0.0111086	0.0061078	0.743
L^2 error	0.0352813	0.0235541	0.0145400	0.0079232	0.716
L^∞ error	0.0819148	0.0502886	0.0331277	0.0195358	0.681

Parameters 4.4. Take $a = 2(.5 + .4\sin(4\pi x))$, $b = (.5 + .4\sin(4\pi x))$, and $\theta(x) = \sin(2\pi x)$. The domain is $x \in [-.5, .5]$. The initial conditions are $\mathbf{u}_0(x) = (e^{-40\tan(\pi x)^2}, 0, 0)$, $S_0(x) = x$. Let $\varepsilon = 1/(500\pi)$ and $\omega = 1/20$. The final time is $t_f = .05$. The step size is $h = 1/400$.

Remark 4.1. Even though this computation is Eulerian, we need not perform a simulation on phase space because of the simplifications presented in section 3.1. Thus the domain listed in the above parameters is only one-dimensional.

The result is illustrated in Figure 6 and shows good agreement.

4.2. Convergence tests. To check convergence of the Lagrangian Gaussian beams, we return to the one-dimensional problem of section 4.1 along with the following.

Parameters 4.5. Take $a = 2(.5 + .4\sin(4\pi x))$, $b = (.5 + .4\sin(4\pi x))$, and $\theta(x) = \sin(2\pi x)$. The domain is $x \in [-.5, .5]$. The initial conditions are $u_0(x) = e^{-40\tan(\pi x)^2}$, $S_0(x) = x$. Let $\omega = 1/15$. The final time is $t_f = .05$.

The values for ε, the number of beams, and the errors in the L^1, L^2, and L^∞ norms for each numerical run are shown in Table 1. The reference solutions were

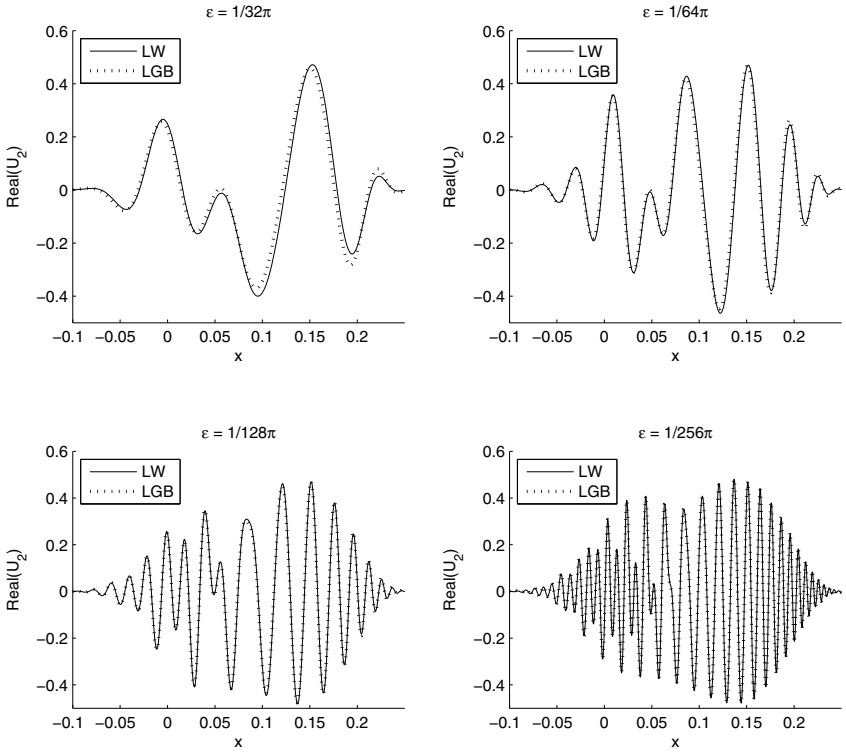

FIG. 7. *Convergence of Lagrangian Gaussian beam method for Parameters* 4.5. *The plots show the real part of the second component of the solution for successively smaller* ε. *The solid line was computed using a converged Lax–Wendroff scheme. The dotted line was computed using Lagrangian Gaussian beams. The errors are reported in Table* 1.

TABLE 2
Convergence data for Eulerian Gaussian beam method.

ε	$\frac{1}{16\pi}$	$\frac{1}{32\pi}$	$\frac{1}{64\pi}$	$\frac{1}{128\pi}$	
Beam count	300	400	550	800	Convergence rate
L^1 error	0.1302495	0.0650149	0.0476501	0.0323120	0.648
L^2 error	0.1423553	0.0743704	0.0557626	0.0378547	0.615
L^∞ error	0.2793297	0.2037504	0.1317435	0.0814930	0.596

computed using a large number of points in the Lax–Wendroff solver. Note that the number of beams used was taken to be proportional to $\varepsilon^{-1/2}$. Figure 7 shows the converging sequence of Lagrangian Gaussian beam solutions by displaying one component of the solution for each ε used in Table 1.

To check convergence of the Eulerian Gaussian beams, a second test was performed with the same parameters as listed above. The results are recorded in Table 2. As can be seen from Tables 1 and 2, the convergence rates are all larger than .5, but none reach or exceed 1. This is expected because Corollary 2.9 implies that convergence should be at least $\mathcal{O}\left(\varepsilon^{1/2}\right)$, as indicated by (2.55).

Remark 4.2. Although not explicitly indicated in Parameters 4.5, Δx and Δt for

these numerical tests were chosen so that the Gaussian beam method converged for each chosen value for ε. This was done to ensure that the convergence rates shown in Tables 1 and 2 are strictly in terms of the decreasing ε and nothing else.

4.3. Two-dimensional Lagrangian. Next consider a nonstrict hyperbolic two-dimensional system. This test is important since the one-dimensional case has many simplifications that do not hold in general for higher-dimensional systems (see section 3.1). In particular, in one dimension, the Hessian matrix M term does not appear in the matrix E^τ given by (2.20). Thus a two-dimensional test is essential.

In reference to (1.1) define

$$A = \begin{pmatrix} 1/a & 0 & 0 & 0 \\ 0 & 1/a & 0 & 0 \\ 0 & 0 & 1/b & 0 \\ 0 & 0 & 0 & 1/b \end{pmatrix}, \quad D^1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

$$D^2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \tag{4.5}$$

where $a, b > 0$ are functions of $\mathbf{x}$. The dispersion matrix is

$$L(\mathbf{x}, \mathbf{k}) = \begin{pmatrix} 0 & 0 & ak_1 & ak_2 \\ 0 & 0 & -ak_2 & ak_1 \\ bk_1 & -bk_2 & 0 & 0 \\ bk_2 & bk_1 & 0 & 0 \end{pmatrix}. \tag{4.6}$$

The eigenvalues and eigenvectors of the dispersion matrix orthonormal with respect to $\langle \cdot, \cdot \rangle_A$ are

$$\begin{cases} H_1(\mathbf{x}, \mathbf{k}) = -\sqrt{ab}\sqrt{k_1^2 + k_2^2}, & \mathbf{b}^{1,1}(\mathbf{x}, \mathbf{k}) = \left(-\sqrt{\frac{a}{2}} \frac{k_2}{\sqrt{k_1^2+k_2^2}}, -\sqrt{\frac{a}{2}} \frac{k_1}{\sqrt{k_1^2+k_2^2}}, 0, \sqrt{\frac{b}{2}} \right), \\ & \mathbf{b}^{1,2}(\mathbf{x}, \mathbf{k}) = \left(-\sqrt{\frac{a}{2}} \frac{k_1}{\sqrt{k_1^2+k_2^2}}, \sqrt{\frac{a}{2}} \frac{k_2}{\sqrt{k_1^2+k_2^2}}, \sqrt{\frac{b}{2}}, 0 \right), \\ H_2(\mathbf{x}, \mathbf{k}) = \sqrt{ab}\sqrt{k_1^2 + k_2^2}, & \mathbf{b}^{2,1}(\mathbf{x}, \mathbf{k}) = \left(\sqrt{\frac{a}{2}} \frac{k_2}{\sqrt{k_1^2+k_2^2}}, \sqrt{\frac{a}{2}} \frac{k_1}{\sqrt{k_1^2+k_2^2}}, 0, \sqrt{\frac{b}{2}} \right), \\ & \mathbf{b}^{2,2}(\mathbf{x}, \mathbf{k}) = \left(\sqrt{\frac{a}{2}} \frac{k_1}{\sqrt{k_1^2+k_2^2}}, -\sqrt{\frac{a}{2}} \frac{k_2}{\sqrt{k_1^2+k_2^2}}, \sqrt{\frac{b}{2}}, 0 \right). \end{cases} \tag{4.7}$$

The coupling matrix for H_1 is

$$N^1 = \begin{pmatrix} 0 & \frac{1}{2}\sqrt{\frac{b}{a}} \frac{1}{\sqrt{k_1^2+k_2^2}} \left((\partial_{x_1} a) k_2 - (\partial_{x_2} a) k_1 \right) \\ -\frac{1}{2}\sqrt{\frac{b}{a}} \frac{1}{\sqrt{k_1^2+k_2^2}} \left((\partial_{x_1} a) k_2 - (\partial_{x_2} a) k_1 \right) & 0 \end{pmatrix} \tag{4.8}$$

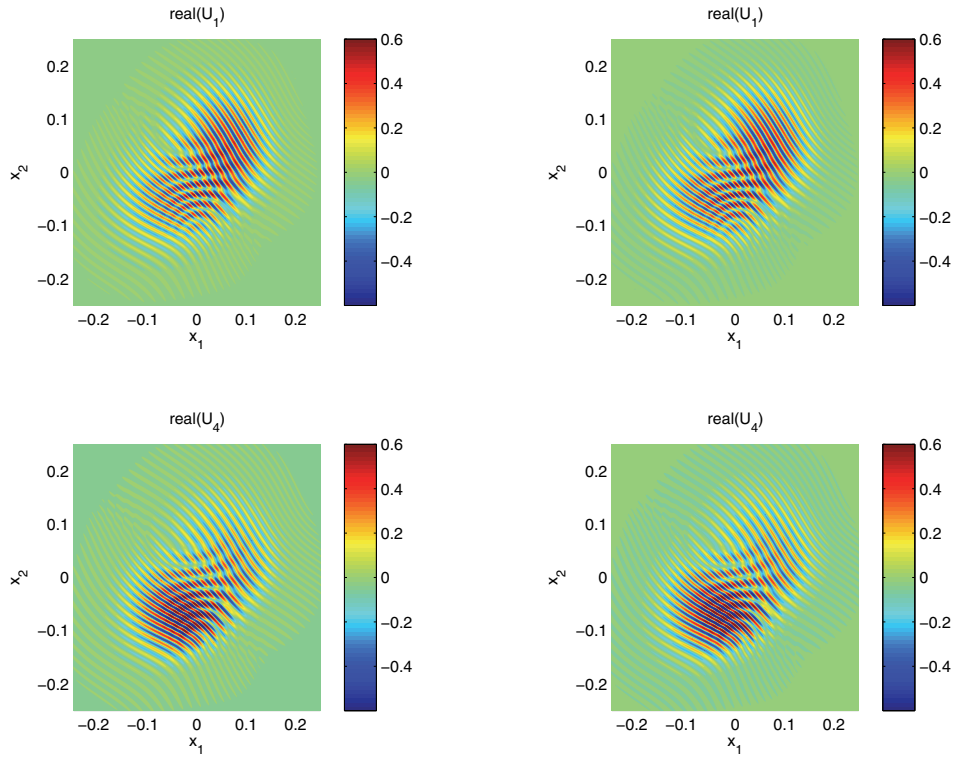

FIG. 8. *Comparison of the real part of the first and fourth components of the Lagrangian Gaussian beam solution for Parameters* 4.6. *The left column is the reference solution, and the right column is the Lagrangian Gaussian beam solution.*

and $N^2 = -N^1$.

For our simulation we take the following.

Parameters 4.6. Take $a(x_1, x_2) = .5 + .4\cos(4\pi x_1)\sin(4\pi x_2)$, $b(x_1, x_2) = 1$. The domain is $(x_1, x_2) \in [-.5, .5] \times [-.5, .5]$ periodic in x_1 and x_2. The initial conditions are $\mathbf{u}_0(x_1, x_2) = (e^{-10(\tan(\pi x_1)^2 + \tan(\pi x_2)^2)}, 0, 0, 0)$, $S_0(x_1, x_2) = x_1 + x_2$. Let $\varepsilon = 1/(100\pi)$ and $\omega = 1/20$. The final time is $t_f = .1$. The number of beams is 60×60 concentrated near the initial conditions.

The result is plotted in Figure 8 and is compared to a reference solution computed using the Lax–Wendroff scheme.

5. Conclusion. We have introduced a Gaussian beam method which solves the high frequency solutions to the linear hyperbolic system (1.1) in the nonstrict hyperbolic case, provided that the system is spectrum preserving in the sense of Definition 2.2. In addition, we have provided convergence results for symmetric hyperbolic systems as well as an Eulerian summation formula that preserves solution accuracy even after the formulation of caustics. Finally, we have provided numerical verification of all methods and formulations introduced. In addition, our derivation for the coupling which occurs in symmetric hyperbolic systems between Gaussian beams within a higher-dimensional eigenspace of the dispersion matrix lays the groundwork for developing Gaussian beam methods for the case where eigenvalues have nonconstant

multiplicity over the domain of computation. This, in turn, leads the way toward developing Gaussian beam methods for other equations with nonconstant eigenvalue multiplicity as well. Nonconstant eigenvalue multiplicity will be the subject of a future paper. As a final note, we point out that within this paper we have provided only an asymptotic justification that our Gaussian beam method converges to the exact solution of (1.1). An elegant rigorous proof of convergence, however, is possible and will appear shortly in a paper to follow.

Appendix A. Gaussian beams for three fundamental examples. The critical step in our derived Gaussian beam method is in assuming the following form for $\mathbf{a}_0$ in our Gaussian beam expansion:

$$d_t T + H_\tau(\mathbf{x}, \tilde{\mathbf{p}}) = 0 \quad \text{and} \quad \mathbf{a}_0 = \sum_{s=1}^{r} c_s(\mathbf{q}, t)\mathbf{b}^{\tau,s}(\mathbf{x}, \tilde{\mathbf{p}}). \tag{A.1}$$

Since $\tilde{\mathbf{p}}$ can be complex, it is critical that we are guaranteed a linearly independent set of eigenvectors $\mathbf{b}^{\tau,s}(\mathbf{x}, \tilde{\mathbf{p}})$ to use in such an expansion. Furthermore, in the subsequent steps in the derivation of our Gaussian beam method which involve taking derivatives of $H_\tau(\mathbf{x}, \tilde{\mathbf{p}})$ and $\mathbf{b}^{\tau,s}(\mathbf{x}, \tilde{\mathbf{p}})$, one must verify that these functions are well defined for complex $\tilde{\mathbf{p}}$ so that using the chain rule is justified. In particular, we need to show that for a given real point $\mathbf{p}$, there exists some complex neighborhood around $\mathbf{p}$ wherein both $H_\tau(\mathbf{x}, \tilde{\mathbf{p}})$ and $\mathbf{b}^{\tau,s}(\mathbf{x}, \tilde{\mathbf{p}})$ are holomorphic.

Since we cannot prove, in general, that the above required details hold, one must check them for each particular symmetric hyperbolic system being considered. To establish the usefulness of our method, we present verification for the three important three-dimensional physical equations considered in [28], namely the acoustic equations, Maxwell's equations, and the elastic wave equations. To this end we begin with some preliminaries.

A.1. Preliminaries. Central to the work to follow is a complete understanding of the matrix

$$P(\mathbf{k}) = \begin{pmatrix} 0 & -k_3 & k_2 \\ k_3 & 0 & -k_1 \\ -k_2 & k_1 & 0 \end{pmatrix}. \tag{A.2}$$

The three eigenvalues of (A.2) may be calculated as $\lambda = 0, \pm\mathbf{i}\sqrt{k_1^2 + k_2^2 + k_3^2}$. Define

$$\zeta(\mathbf{k}) = \sqrt{k_1^2 + k_2^2 + k_3^2} = \sqrt{\mathbf{k} \cdot \mathbf{k}}, \tag{A.3}$$

where we note that $\zeta(\mathbf{k})$ may be complex if $\mathbf{k}$ is complex. In this form, the three eigenvalues are written as $\lambda = 0, \pm\mathbf{i}\zeta$. We now prove a lemma that establishes that in some neighborhood in $\mathbb{C}^3$ about a real point $\mathbf{k}_0 \neq 0$, $\zeta(\mathbf{k})$ defined by (A.3) is holomorphic.

LEMMA A.1. *Suppose $\mathbf{k}_0 \in \mathbb{R}^d$, where $|\mathbf{k}_0| > 0$; then the function $\zeta(\mathbf{k}) = \sqrt{\mathbf{k} \cdot \mathbf{k}}$ is holomorphic inside the ball $B_\delta(\mathbf{k}_0) = \left\{\mathbf{k} \in \mathbb{C}^d : |\mathbf{k} - \mathbf{k}_0| < \delta\right\}$, provided that $\delta < (-1 + \sqrt{2})|\mathbf{k}_0|$.*

Proof. We will show that we may choose an appropriate branch cut for the square root function. We begin by starting with an arbitrary element of the ball $B_\delta(\mathbf{k}_0)$ represented by $\tilde{\mathbf{k}} = \mathbf{k}_0 + \delta\mathbf{k}$, where $|\mathbf{k}| = 1$. Then

$$\tilde{\mathbf{k}} \cdot \tilde{\mathbf{k}} = |\mathbf{k}_0|^2 + 2\delta\mathbf{k}_0 \cdot \mathbf{k} + \delta^2\mathbf{k} \cdot \mathbf{k}. \tag{A.4}$$

Next

$$\begin{aligned}\mathrm{Re}(\tilde{\mathbf{k}}\cdot\tilde{\mathbf{k}}) &= |\mathbf{k}_0|^2 + 2\delta\mathbf{k}_0\cdot\mathrm{Re}(\mathbf{k}) + \delta^2(|\mathrm{Re}(\mathbf{k})|^2 - |\mathrm{Im}(\mathbf{k})|^2)\\ &\geq |\mathbf{k}_0|^2 - 2\delta|\mathbf{k}_0| - \delta^2\\ &> 0,\end{aligned} \tag{A.5}$$

where the final inequality follows from $\delta < (-1+\sqrt{2})|\mathbf{k}_0|$. Thus since $\mathrm{Re}(\tilde{\mathbf{k}}\cdot\tilde{\mathbf{k}}) > 0$, we may choose the branch cut of the square root part of $\zeta(\mathbf{k})$ to be along the negative real axis. Thus on $B_\delta(\mathbf{k}_0)$, $\zeta(\mathbf{k}) = \sqrt{\mathbf{k}\cdot\mathbf{k}}$ is a composition of holomorphic functions and so is itself holomorphic. □

Next we establish that in some neighborhood in $\mathbb{C}^3$ about a real point $\mathbf{k}_0 \neq \mathbf{0}$, the eigenvectors of (A.2) are linearly independent and holomorphic. To do this, we simply present the eigenvectors as follows.

For $\lambda = 0$ the holomorphic eigenvector is

$$\mathbf{v}_0(\mathbf{k}) = \hat{\mathbf{k}} \equiv \mathbf{k}/\zeta(\mathbf{k}). \tag{A.6}$$

For $\lambda = \pm\mathbf{i}\zeta(\mathbf{k})$ we write the holomorphic eigenvectors in two separate regimes. First, assuming that $k_1 \neq 0$ or $k_2 \neq 0$, we may write

$$\mathbf{v}_\pm(\mathbf{k}) = \frac{1}{\zeta(\mathbf{k})\sqrt{2\,(k_1^2+k_2^2)}}\left(-k_1k_3 \mp \mathbf{i}k_2\zeta(\mathbf{k}), -k_2k_3 \pm \mathbf{i}k_1\zeta(\mathbf{k}), k_1^2+k_2^2\right). \tag{A.7}$$

Second, assuming that $k_1 \neq 0$ or $k_3 \neq 0$, we may write

$$\mathbf{v}_\pm(\mathbf{k}) = \frac{1}{\zeta(\mathbf{k})\sqrt{2\,(k_1^2+k_3^2)}}\left(-k_1k_2 \pm \mathbf{i}k_3\zeta(\mathbf{k}), k_1^2+k_3^2, -k_2k_3 \mp \mathbf{i}k_1\zeta(\mathbf{k})\right). \tag{A.8}$$

We remark that in the neighborhood $B_\delta(\mathbf{k}_0)$ of any real point $\mathbf{k}_0 \neq \mathbf{0}$ defined by Lemma A.1, the vector $\mathbf{v}_0(\mathbf{k})$ is a holomorphic function of $\mathbf{k}$. By again using Lemma A.1 with $d = 2$, we may conclude that $\mathbf{v}_\pm(\mathbf{k})$ given by (A.7) are holomorphic inside $B_\delta(\mathbf{k}_0)$ when $\delta < (-1+\sqrt{2})\sqrt{k_{0,1}^2+k_{0,2}^2}$, and $\mathbf{v}_\pm(\mathbf{k})$ given by (A.8) are holomorphic inside $B_\delta(\mathbf{k}_0)$ when $\delta < (-1+\sqrt{2})\sqrt{k_{0,1}^2+k_{0,3}^2}$. Finally, we observe that all three eigenvectors are linearly independent, provided that $\mathbf{k}$ remains within $B_\delta(\mathbf{k}_0)$ simply because the eigenvalues are all distinct in this neighborhood. Note that when $\mathbf{k}_0 \neq 0$ is real, then the vectors $\mathbf{v}_0(\mathbf{k}_0), \mathbf{v}_\pm(\mathbf{k}_0)$ are orthonormal. However, orthonormality does not hold inside of the complex ball $B_\delta(\mathbf{k}_0)$ no matter which δ is chosen.

We now have a relatively complete understanding of the eigenstructure of the matrix (A.2), so we finish this section by introducing two new vectors which will become useful in the following sections:

$$\mathbf{z}_1(\mathbf{k}) = \frac{1}{\sqrt{2}}\left(\mathbf{v}_+(\mathbf{k}) + \mathbf{v}_-(\mathbf{k})\right) \quad \text{and} \quad \mathbf{z}_2(\mathbf{k}) = \frac{\mathbf{i}}{\sqrt{2}}\left(\mathbf{v}_+(\mathbf{k}) - \mathbf{v}_-(\mathbf{k})\right). \tag{A.9}$$

One may easily show that for these vectors,

$$P(\mathbf{k})\mathbf{z}_1(\mathbf{k}) = \zeta(\mathbf{k})\mathbf{z}_2(\mathbf{k}) \quad \text{and} \quad P(\mathbf{k})\mathbf{z}_2(\mathbf{k}) = -\zeta(\mathbf{k})\mathbf{z}_1(\mathbf{k}), \tag{A.10}$$

and also that they are linearly independent, provided that $\mathbf{v}_\pm(\mathbf{k})$ are linearly independent. Finally, note that when $\mathbf{k}$ is real, $\mathbf{z}_1(\mathbf{k})$ and $\mathbf{z}_2(\mathbf{k})$ in (A.9) are also orthonormal.

A.2. Acoustic equations. Referring the reader to [28] on the setup for the acoustic equations, we summarize as follows. The dispersion matrix in block form is given by

$$L = \begin{pmatrix} 0 & \mathbf{k}/\rho \\ \mathbf{k}^T/\kappa & 0 \end{pmatrix}, \tag{A.11}$$

where $\rho(\mathbf{x})$ is the density and $\kappa(\mathbf{x})$ is the compressibility. The eigenvalues of the dispersion matrix are $\lambda_0 = 0$ with multiplicity two and $\lambda_\pm = \pm v(\mathbf{x})\zeta(\mathbf{k})$ with multiplicity one. The eigenvectors are

$$\begin{aligned} \mathbf{b}^{(0,1)}(\mathbf{x},\mathbf{k}) &= \left(\mathbf{z}_1(\mathbf{k})/\sqrt{\rho}, 0\right), \\ \mathbf{b}^{(0,2)}(\mathbf{x},\mathbf{k}) &= \left(\mathbf{z}_2(\mathbf{k})/\sqrt{\rho}, 0\right), \\ \mathbf{b}^{+}(\mathbf{x},\mathbf{k}) &= \left(\hat{\mathbf{k}}/\sqrt{2\rho}, 1/\sqrt{2\kappa}\right), \\ \mathbf{b}^{-}(\mathbf{x},\mathbf{k}) &= \left(\hat{\mathbf{k}}/\sqrt{2\rho}, -1/\sqrt{2\kappa}\right). \end{aligned} \tag{A.12}$$

From these explicit forms for the eigenvalues and eigenvectors, it is clear that they are holomorphic for $\mathbf{k} \in B_\delta(\mathbf{k}_0)$ for any real point $\mathbf{k}_0 \neq 0$ (as discussed in section A.1). Linear independence of the eigenvectors inside $B_\delta(\mathbf{k}_0)$ follows from the linear independence of $\hat{\mathbf{k}}$, $\mathbf{z}_1$, and $\mathbf{z}_2$. Finally, note that (A.12) are orthonormal in $\langle \cdot, \cdot \rangle_A$ when $\mathbf{k}$ is real.

A.3. Maxwell's equations. Referring the reader to [28] on the setup for Maxwell's equations, we summarize as follows. The dispersion matrix for this problem is given by

$$L = \begin{pmatrix} 0 & -\frac{1}{\epsilon}P(\mathbf{k}) \\ \frac{1}{\mu}P(\mathbf{k}) & 0 \end{pmatrix}, \tag{A.13}$$

where $P(\mathbf{k})$ is given by (A.2), $\epsilon(\mathbf{x})$ is the dielectric permittivity, and $\mu(\mathbf{x})$ is the relative magnetic permeability. We will use $\mathbf{z}_1$ and $\mathbf{z}_2$, defined by (A.9), to construct the eigenvectors of the dispersion matrix (A.13). The three eigenvalues of L, each with multiplicity two, are given by $\lambda_0 = 0$, $\lambda_+(\mathbf{x},\mathbf{k}) = v(\mathbf{x})\zeta(\mathbf{k})$, $\lambda_-(\mathbf{x},\mathbf{k}) = -v(\mathbf{x})\zeta(\mathbf{k})$ with

$$v(\mathbf{x}) = \frac{1}{\sqrt{\epsilon(\mathbf{x})\mu(\mathbf{x})}} \tag{A.14}$$

and $\zeta(\mathbf{k})$ defined by (A.9). Proving that these are the eigenvalues is a matter of writing down the eigenvectors. For $\lambda_0 = 0$ one has

$$\mathbf{b}^{(0,1)} = \frac{1}{\sqrt{\epsilon}}(\hat{\mathbf{k}}, 0), \quad \mathbf{b}^{(0,2)} = \frac{1}{\sqrt{\mu}}(0, \hat{\mathbf{k}}). \tag{A.15}$$

For $\lambda_+ = v\zeta$ one has

$$\mathbf{b}^{(+,1)} = \left(\sqrt{\frac{1}{2\epsilon}}\mathbf{z}_1, \sqrt{\frac{1}{2\mu}}\mathbf{z}_2\right), \quad \mathbf{b}^{(+,2)} = \left(\sqrt{\frac{1}{2\epsilon}}\mathbf{z}_2, -\sqrt{\frac{1}{2\mu}}\mathbf{z}_1\right). \tag{A.16}$$

For $\lambda_- = -v\zeta$ one has

$$\mathbf{b}^{(-,1)} = \left(\sqrt{\frac{1}{2\epsilon}}\mathbf{z}_1, -\sqrt{\frac{1}{2\mu}}\mathbf{z}_2\right), \quad \mathbf{b}^{(-,2)} = \left(\sqrt{\frac{1}{2\epsilon}}\mathbf{z}_2, \sqrt{\frac{1}{2\mu}}\mathbf{z}_1\right). \tag{A.17}$$

We note that all of these eigenvectors will be linearly independent, provided that $\mathbf{z}_\pm$ are linearly independent. However, when $\mathbf{k}$ is complex, these vectors may not be orthogonal. Finally, note that these eigenvectors are orthonormal in $\langle\cdot,\cdot\rangle_A$ when $\mathbf{k}$ is real.

A.4. Elastic equations. Referring the reader to [28] on the setup for the elastic equations, we summarize as follows. The dispersion matrix for this problem is given in block form by

$$L = -\begin{pmatrix} 0 & K(\mathbf{k})/\rho & M(\mathbf{k})/\rho & \mathbf{k}/\rho \\ 2\mu K(\mathbf{k}) & 0 & 0 & 0 \\ \mu M(\mathbf{k}) & 0 & 0 & 0 \\ \lambda\mathbf{k}^T & 0 & 0 & 0 \end{pmatrix}, \tag{A.18}$$

where $K(\mathbf{k}) = \operatorname{diag}(k_1, k_2, k_3)$ and

$$M(\mathbf{k}) = \begin{pmatrix} 0 & k_3 & k_2 \\ k_3 & 0 & k_1 \\ k_2 & k_1 & 0 \end{pmatrix}. \tag{A.19}$$

By writing down the appropriate eigenvectors we will show that the eigenvalues of the above dispersion matrix for $\mathbf{k} \in B_\delta(\mathbf{k}_0)$ about a real point $\mathbf{k}_0 \neq 0$ (as discussed in section A.1) are given by

$$\begin{aligned} &\lambda_0 = 0 \quad \text{with multiplicity four,} \\ &\lambda_\pm^P = \pm v_P(\mathbf{x})\zeta(\mathbf{k}) \quad \text{with multiplicity one, and} \\ &\lambda_\pm^S = \pm v_S(\mathbf{x})\zeta(\mathbf{k}) \quad \text{with multiplicity two.} \end{aligned} \tag{A.20}$$

Note that the velocities are given by

$$v_P(\mathbf{x}) = \sqrt{(2\mu(\mathbf{x}) + \lambda(\mathbf{x}))/\rho(\mathbf{x})}, \quad v_S(\mathbf{x}) = \sqrt{\mu(\mathbf{x})/\rho(\mathbf{x})}, \tag{A.21}$$

where $\mu(\mathbf{x})$ and $\lambda(\mathbf{x})$ are the Lame parameters and $\rho(\mathbf{x})$ is the density (see [28] for details). Then the eigenvectors are

(A.22)

$$\begin{aligned} &\mathbf{b}_\pm^P(\mathbf{x},\mathbf{k}) \\ &\quad = \left(\mathbf{k}/\sqrt{2\rho}, \mp 2\mu K(\mathbf{k})\mathbf{k}/\sqrt{2(2\mu+\lambda)}, \mp\mu M(\mathbf{k})\mathbf{k}/\sqrt{2(2\mu+\lambda)}, \mp\lambda/\sqrt{2(2\mu+\lambda)}\right), \\ &\mathbf{b}_\pm^{S_j}(\mathbf{x},\mathbf{k}) = \left(\mathbf{z}_j/\sqrt{2\rho}, \mp 2\sqrt{\mu/2}K(\mathbf{k})\mathbf{z}_j, \mp\sqrt{\mu/2}M(\mathbf{k})\mathbf{z}_j, 0\right), \quad j = 1,2, \\ &\mathbf{b}^{(0,j)}(\mathbf{x},\mathbf{k}) = \left(0, \sqrt{2\mu}K(\mathbf{z}_j)\mathbf{z}_j, \sqrt{\mu/2}M(\mathbf{z}_j)\mathbf{z}_j, 0\right), \quad j = 1,2, \\ &\mathbf{b}^{(0,3)}(\mathbf{x},\mathbf{k}) = \left(0, 2\sqrt{\mu}K(\mathbf{z}_1)\mathbf{z}_2, \sqrt{\mu}M(\mathbf{z}_1)\mathbf{z}_2, 0\right), \\ &\mathbf{b}^{(0,4)}(\mathbf{x},\mathbf{k}) \\ &\quad = \left(0, 2\sqrt{\lambda\mu}K(\hat{\mathbf{k}})\hat{\mathbf{k}}/\sqrt{2(2\mu+\lambda)}, \sqrt{\lambda\mu}M(\hat{\mathbf{k}})\hat{\mathbf{k}}/\sqrt{2(2\mu+\lambda)}, -2\sqrt{\lambda\mu}/\sqrt{2(2\mu+\lambda)}\right), \end{aligned}$$

where $\mathbf{z}_1$ and $\mathbf{z}_2$ are given by (A.9) and $\hat{\mathbf{k}}$ is defined by (A.6). That equations (A.22) are truly the eigenvectors relies on a couple of facts: First,

$$2K(\mathbf{k})^2 + M(\mathbf{k})^2 = \mathbf{k}\mathbf{k}^T + \mathbf{k}\cdot\mathbf{k}I. \tag{A.23}$$

Second,

$$\begin{aligned} \mathbf{k}\cdot\mathbf{z}_1 &= -(\mathbf{k}^T P(\mathbf{k}))\mathbf{z}_2/\zeta(\mathbf{k}) = 0, \\ \mathbf{k}\cdot\mathbf{z}_2 &= (\mathbf{k}^T P(\mathbf{k}))\mathbf{z}_1/\zeta(\mathbf{k}) = 0, \end{aligned} \tag{A.24}$$

provided that $\zeta(\mathbf{k}) \neq 0$. That functions in (A.22) are holomorphic functions of $\mathbf{k}$ inside some neighborhood $B_\delta(\mathbf{k}_0)$ of a real point $\mathbf{k}_0 \neq 0$ follows from Lemma A.1, and the linear independence of these vectors inside $B_\delta(\mathbf{k}_0)$ follows from the linear independence of $\mathbf{k}$, $\mathbf{z}_1$, and $\mathbf{z}_2$. Finally, note that the eigenvectors given by (A.22) are orthonormal in $\langle\cdot,\cdot\rangle_A$ when $\mathbf{k}$ is real.

Appendix B. Simplification of the matrix E^τ. The purpose of this appendix is to justify the simplification of the matrix E^τ given by (2.20) to the form shown by (2.23). Since τ does not affect the following computation in any way, it will be dropped from the notation for the sake of clean presentation.

Begin by restating (2.20):

$$E_{is} = \langle A\mathbf{b}^i, A^{-1}D^j\{\nabla_{\mathbf{q}}\mathbf{b}^s + (\nabla_{\mathbf{p}}\mathbf{b}^s)M\}\mathbf{e}^j - \nabla_{\mathbf{p}}\mathbf{b}^s[\nabla_{\mathbf{q}}H + M\nabla_{\mathbf{p}}H]\rangle. \tag{B.1}$$

Consider for now only the terms involving the matrix $M(t,\mathbf{q})$, which, we recall, is symmetric (see Theorem 2.6). In particular, we wish to rewrite the expression

$$A^{-1}D^j(\nabla_{\mathbf{p}}\mathbf{b}^s)M\mathbf{e}^j - \nabla_{\mathbf{p}}\mathbf{b}^s M\nabla_{\mathbf{p}}H. \tag{B.2}$$

Using Einstein summation notation, begin with the statement that the vector $\mathbf{b}^s$ is in the null space of $A^{-1}D^j k_j - HI$ which reads as

$$\left[A^{-1}_{ml}D^n_{lj}p_n - H\delta_{mj}\right]b^s_j = 0. \tag{B.3}$$

Taking the partial derivative with respect to $\mathbf{p}$ gives

$$\left[A^{-1}_{ml}D^f_{lj} - \left(\partial_{p_f}H\right)\delta_{mj}\right]b^s_j + \left[A^{-1}_{ml}D^n_{lj}p_n - H\delta_{mj}\right]\left(\partial_{p_f}b^s_j\right) = 0. \tag{B.4}$$

Multiplying by the matrix M gives

$$\left[A^{-1}_{ml}D^f_{lj}M_{fg} - \left(\partial_{p_f}HM_{fg}\right)\delta_{mj}\right]b^s_j + \left[A^{-1}_{ml}D^n_{lj}p_n - H\delta_{mj}\right]\left(\partial_{p_f}b^s_jM_{fg}\right) = 0. \tag{B.5}$$

Taking the partial with respect to $\mathbf{p}$ again and summing over the index g gives

$$\begin{aligned} &-\left(\partial_{p_f}\partial_{p_g}H\right)M_{fg}\delta_{mj}b^s_j \\ &+A^{-1}_{ml}D^f_{lj}M_{fg}\left(\partial_{p_g}b^s_j\right) + A^{-1}_{ml}D^g_{lj}M_{fg}\left(\partial_{p_f}b^s_j\right) \\ &-\left(\partial_{p_f}H\right)M_{fg}\delta_{mj}\left(\partial_{p_g}b^s_j\right) - \left(\partial_{p_g}H\right)M_{fg}\delta_{mj}\left(\partial_{p_f}b^s_j\right) \\ &+\left[A^{-1}_{ml}D^n_{lj}p_n - H\delta_{mj}\right]\left(M_{fg}\partial_{p_g}\partial_{p_f}b^s_j\right) = 0. \end{aligned} \tag{B.6}$$

Because of the symmetry of M the two terms on the second line of (B.6) are the same, and the two terms on the third line of (B.6) are also the same. Thus (B.6) may be written as

(B.7)
$$\begin{aligned} &A^{-1}_{ml}D^g_{lj}M_{fg}\left(\partial_{p_f}b^s_j\right) - \left(\partial_{p_f}H\right)M_{fg}\delta_{mj}\left(\partial_{p_g}b^s_j\right) \\ &\qquad = \tfrac{1}{2}\left\{\left(\partial_{p_f}\partial_{p_g}H\right)M_{fg}\delta_{mj}b^s_j - \left[A^{-1}_{ml}D^n_{lj}p_n - H\delta_{mj}\right]\left(M_{fg}\partial_{p_g}\partial_{p_f}b^s_j\right)\right\}. \end{aligned}$$

Note that the expression on the left side of (B.7) is exactly (B.2). Additionally recall that the dispersion matrix is self-adjoint in $\langle\cdot,\cdot\rangle_A$ so that the term $A^{-1}_{ml}D^n_{lj}p_n - H\delta_{mj}$ is also self-adjoint in $\langle\cdot,\cdot\rangle_A$. Using these facts, one has

(B.8)
$$\begin{aligned} &\langle A\mathbf{b}^i, A^{-1}D^j(\nabla_{\mathbf{p}}\mathbf{b}^s)M\mathbf{e}^j - \nabla_{\mathbf{p}}\mathbf{b}^s M\nabla_{\mathbf{p}}H\rangle \\ &\quad = A_{mh}b^i_h\tfrac{1}{2}\left\{\left(\partial_{p_f}\partial_{p_g}H\right)M_{fg}\delta_{mj}b^s_j - \left[A^{-1}_{ml}D^n_{lj}p_n - H\delta_{mj}\right]\left(M_{fg}\partial_{p_g}\partial_{p_f}b^s_j\right)\right\} \\ &\quad = \tfrac{1}{2}\left[\left(\partial_{p_f}\partial_{p_g}H\right)M_{fg}\right]\delta_{is} - \tfrac{1}{2}\left[A^{-1}_{ml}D^n_{lj}p_n - H\delta_{mj}\right]b^i_jA_{mh}\left(M_{fg}\partial_{p_g}\partial_{p_f}b^s_h\right) \\ &\quad = \tfrac{1}{2}\left[\left(\partial_{p_f}\partial_{p_g}H\right)M_{fg}\right]\delta_{is} \\ &\quad = \tfrac{1}{2}\mathrm{Tr}\left[M\nabla_{\mathbf{pp}}H\right]\delta_{is}. \end{aligned}$$

This result shows that the off-diagonal terms (B.1) are not influenced by the matrix M.

Next derive the contribution to the diagonal of (B.1) from the terms not involving M. One gets

(B.9)
$$\begin{aligned} \langle A\mathbf{b}^s, A^{-1}D^j\nabla_{\mathbf{q}}\mathbf{b}^s\mathbf{e}^j - \nabla_{\mathbf{p}}\mathbf{b}^s\nabla_{\mathbf{q}}H\rangle &= \langle\mathbf{b}^s, D^j\nabla_{\mathbf{q}}\mathbf{b}^s\mathbf{e}^j\rangle - \langle A\mathbf{b}^s, \nabla_{\mathbf{p}}\mathbf{b}^s\nabla_{\mathbf{q}}H\rangle \\ &= \langle\mathbf{b}^s, D^j\partial_{q_j}\mathbf{b}^s\rangle - \left(\partial_{q_j}H\right)\langle A\mathbf{b}^s, \partial_{p_j}\mathbf{b}^s\rangle \\ &= \tfrac{1}{2}\partial_{q_j}\langle\mathbf{b}^s, D^j\mathbf{b}^s\rangle - \tfrac{1}{2}\left(\partial_{q_j}H\right)\partial_{p_j}\langle A\mathbf{b}^s, \mathbf{b}^s\rangle \\ &= \tfrac{1}{2}\partial_{q_j}\partial_{p_j}H - \tfrac{1}{2}\left(\partial_{q_j}H\right)\partial_{p_j}1 \\ &= \tfrac{1}{2}\partial_{q_j}\partial_{p_j}H. \end{aligned}$$

In keeping with [28], define the skew symmetric matrix

$$N_{is} = \langle A\mathbf{b}^i, A^{-1}D^j\nabla_{\mathbf{q}}\mathbf{b}^s\mathbf{e}^j - \nabla_{\mathbf{p}}\mathbf{b}^s\nabla_{\mathbf{q}}H\rangle - \frac{1}{2}\nabla_{\mathbf{q}}\cdot\nabla_{\mathbf{p}}H\delta_{is}. \tag{B.10}$$

Finally, (B.1) may be written as

(B.11)
$$\begin{aligned} E_{is} &= \tfrac{1}{2}\left\{\nabla_{\mathbf{q}}\cdot\nabla_{\mathbf{p}}H + \mathrm{Tr}\left[M\nabla_{\mathbf{pp}}H\right]\right\}\delta_{is} + N_{is} \\ &= \tfrac{1}{2}\mathrm{Tr}\left[\nabla_{\mathbf{qp}}H + M\nabla_{\mathbf{pp}}H\right]\delta_{is} + N_{is}. \end{aligned}$$

This is what we set out to show.

REFERENCES

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

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

[3] V. Červený, M. Popov, and I. Pšenčík, *Computation of wave fields in inhomogeneous media—Gaussian beam approach*, Geophys. J. R. Astr. Soc., 70 (1982), pp. 109–128.
[4] L. Chai, S. Jin, and Q. Li, *Semi-classical models for the Schrödinger equation with periodic potentials and band crossings*, Kinet. Relat. Models, 6 (2013), pp. 505–532.
[5] L. Dieci and T. Eirola, *Preserving monotonicity in the numerical solution of Riccati differential equations*, Numer. Math., 74 (1996), pp. 35–47.
[6] B. Engquist and O. Runborg, *Computational high frequency wave propagation*, Acta Numer., 12 (2003), pp. 181–266.
[7] E. Faou, V. Gradinaru, and C. Lubich, *Computing semiclassical quantum dynamics with Hagedorn wavepackets*, SIAM J. Sci. Comput., 31 (2009), pp. 3027–3041.
[8] P. Gérard, P. A. Markowich, N. J. Mauser, and F. Poupaud, *Homogenization limits and Wigner transforms*, Comm. Pure Appl. Math., 50 (1997), pp. 323–379.
[9] G. A. Hagedorn, *Raising and lowering operators for semiclassical wave packets*, Ann. Physics, 269 (1998), pp. 77–104.
[10] E. J. Heller, *Frozen Gaussians: A very simple semiclassical approximation*, J. Chem. Phys., 75 (1981), pp. 2923–2931.
[11] L. Jefferis and S. Jin, *Computing high frequency solutions of symmetric hyperbolic systems with polarized waves*, Commun. Math. Sci., 13 (2015), pp. 1001–1024.
[12] S. Jin, P. Markowich, and C. Sparber, *Mathematical and computational methods for semiclassical Schrödinger equations*, Acta Numer., 20 (2012), pp. 1–89.
[13] S. Jin, P. Qi, and Z. Zhang, *An Eulerian surface hopping method for the Schrödinger equation with conical crossings*, Multiscale Model. Simul., 9 (2011), pp. 258–281.
[14] S. Jin, H. Wu, and X. Yang, *Gaussian beam methods for the Schrödinger equation in the semiclassical regime: Lagrangian and Eulerian formulations*, Commun. Math. Sci., 6 (2008), pp. 995–1020.
[15] S. Jin, H. Wu, and X. Yang, *Semi-Eulerian and high order Gaussian beam methods for the Schrödinger equation in the semiclassical regime*, Commun. Comput. Phys., 9 (2011), pp. 668–687.
[16] S. Jin, H. Wu, X. Yang, and Z. Huang, *Bloch decomposition-based Gaussian beam methods for the Schrödinger equation with periodic potentials*, J. Comput. Phys., 229 (2010), pp. 4869–4883.
[17] L. Klimeš, *Optimization of the shape of Gaussian beams of a fixed length*, Stud. Geoph. et Geod., 33 (1989), pp. 146–163.
[18] C. Lasser, T. Swart, and S. Teufel, *Construction and validation of a rigorous surface hopping algorithm for conical crossings*, Commun. Math. Sci., 5 (2007), pp. 789–814.
[19] S. Leung, J. Qian, and R. Burridge, *Eulerian Gaussian beams for high-frequency wave propagation*, Geophys., 72 (2007), pp. 61–76.
[20] P.-L. Lions and T. Paul, *Sur les mesures de Wigner*, Rev. Mat. Iberoamericana, 9 (1993), pp. 553–618.
[21] H. Liu, O. Runborg, and N. M. Tanushev, *Error estimates for Gaussian beam superpositions*, Math. Comp., 82 (2013), pp. 919–952.
[22] J. Lu and X. Yang, *Frozen Gaussian approximation for high frequency wave propagation*, Commun. Math. Sci., 9 (2011), pp. 663–683.
[23] J. Lu and X. Yang, *Frozen Gaussian approximation for general linear strictly hyperbolic systems: Formulation and Eulerian methods*, Multiscale Model. Simul., 10 (2012), pp. 451–472.
[24] C. Min, *Local level set method in high dimension and codimension*, J. Comput. Phys., 200 (2004), pp. 368–382.
[25] M. Popov, *A new method of computation of wave fields using Gaussian beams*, Wave Motion, 4 (1982), pp. 85–97.
[26] J. Qian and L. Ying, *Fast Gaussian wavepacket transforms and Gaussian beams for the Schrödinger equation*, J. Comput. Phys., 229 (2010), pp. 7848–7873.
[27] J. Ralston, *Gaussian beams and the propagation of singularities*, in Studies in Partial Differential Equations, MAA Stud. Math. 23, Mathematical Association of America, Washington, DC, 1982, pp. 206–248.
[28] L. Ryzhik, G. Papanicolaou, and J. Keller, *Transport equations for elastic and other waves in random media*, Wave Motion, 24 (1996), pp. 327–370.
[29] G. Sundaram and Q. Niu, *Wave-packet dynamics in slowly perturbed crystals: Gradient corrections and Berry-phase effects*, Phys. Rev. B, 59 (1999), pp. 14915–14925.
[30] N. M. Tanushev, *Superpositions and higher order Gaussian beams*, Commun. Math. Sci., 6 (2008), pp. 449–475.
[31] N. M. Tanushev, B. Engquist, and R. Tsai, *Gaussian beam decomposition of high frequency wave fields*, J. Comput. Phys., 228 (2009), pp. 8856–8871.

[32] J. Tully, *Molecular dynamics with electronic transitions*, J. Chem. Phys., 93 (1990), pp. 1061–1071.
[33] B. S. White, A. Norris, A. Bayliss, and R. Burridge, *Some remarks on the Gaussian beam summation method*, Geophys. J. R. Astr. Soc., 89 (1987), pp. 579–636.
[34] E. Wigner, *On the quantum correction for thermodynamic equilibrium*, Phys. Rev., 40 (1932), pp. 749–759.