A Wavefront-Based Gaussian Beam Method for Computing High Frequency Wave Propagation Problems

Mohammad Motamed[a,*], Olof Runborg[b,c]

[a] *Department of Mathematics and Statistics, The University of New Mexico, Albuquerque, NM 87131, USA*
[b] *Department of Mathematics, KTH Royal Institute of Technology, Stockholm, Sweden*
[c] *Swedish e-Science Research Center (SeRC), KTH Royal Institute of Technology, Stockholm, Sweden*

Abstract

We present a novel wavefront method based on Gaussian beams for computing high frequency wave propagation problems. Unlike standard geometrical optics, Gaussian beams compute the correct solution of the wave field also at caustics. The method tracks a front of two canonical beams with two particular initial values for width and curvature. In a fast post-processing step, from the canonical solutions we recreate any other Gaussian beam with arbitrary initial data on the initial front. This provides a simple mechanism to include a variety of optimization processes, including error minimization and beam width minimization, for a posteriori selection of optimal beam initial parameters. The performance of the method is illustrated with two numerical examples.

Keywords: wave propagation, high frequency, asymptotic approximation, summation of Gaussian beams, wavefront methods

1. Introduction

In direct discretization methods for high frequency wave problems, a large number of grid points is needed to resolve the wave oscillations, and the computational cost to maintain constant accuracy grows algebraically with the frequency. At sufficiently high frequencies, direct simulations are not feasible. As an alternative, one can use high frequency asymptotic methods where the cost is either independent of or grows slowly with the frequency, see [1, 2]. The Gaussian beam method is one such asymptotic method for computing high frequency wave fields in smoothly varying inhomogeneous media. It was proposed by Popov [3], based on earlier work of Babic and Pankratova [4]. The method was first applied by Katchalov and Popov [5], Cerveny et al. [6] and Klimeš [7] to describe high-frequency seismic wave fields by the summation of Gaussian beams. In quantum chemistry, Gaussian beams are higher order versions of classical coherent states, and they are used to approximate the Schrödinger equation; see e.g. Heller, Herman and Kluk [8, 9]. Gaussian beams were later applied to

*Corresponding author
Email addresses: motamed@math.unm.edu (Mohammad Motamed), olofr@nada.kth.se (Olof Runborg)

seismic migration by Hill [10, 11]. For a rigorous mathematical analysis of Gaussian beams we refer to [12] and the more recent investigations on accuracy [13, 14, 15, 16, 17, 18]. The main advantage of this method is that Gaussian beams provide the correct solution also at caustics where standard geometrical optics breaks down.

In the Gaussian beam method, the initial/boundary data or the wave sources which generate the high frequency wave field are decomposed into Gaussian beams. Individual Gaussian beams are computed in a Lagrangian fashion by ray tracing, where quantities such as the curvature and width of beams are calculated from ordinary differential equations (ODEs) along the central ray of the beams. The initial conditions for the ODEs are obtained from the field decomposition at the boundary or the source. The contributions of the beams concentrated close to their central rays are determined by Taylor expansion. The wave field at a receiver is then obtained as a weighted superposition of the Gaussian beams situated close to the receiver.

The past few years have seen a renewed interest in Gaussian beam based methods and their applications [19, 20, 21, 22]. One new direction is the Eulerian Gaussian beam summation methods [23, 24, 25, 26]. In this approach, the problem is formulated by Liouville-type equations in phase space giving uniformly distributed Eulerian traveltimes and amplitudes for multiple sources. A recent survey of Gaussian beam methods can be found in [27]. Numerical approaches for treating general high frequency initial data for superposition over physical space were considered in [28, 29] for the wave equation.

In this paper, we revisit the Lagrangian formulation and present a wavefront method for computing Gaussian beams. Wavefront methods have been very successful for standard geometrical optics as they provide a simple mechanism for controlling the resolution and accuracy of the numerical approximation [30, 31]. Using them with Gaussian beams is not as straightforward since the beam method strongly depends on the distribution and width of the beams at the initial front and on how they spread during their evolution, see e.g. [17, 32, 33]. We construct our novel wavefront method based on two canonical functions. We present an efficient strategy consisting of two parts: (1) We compute the wavefronts together with a set of canonical solutions with a priori and fixed initial data; and (2) In a post-processing step, from the canonical solutions we recreate Gaussian beams with a posteriori, optimal selection of initial data and compute the wave field by a weighted sum of beams. This strategy has a few advantages. First, we can compute beams with any arbitrary initial conditions by a simple linear combination of the canonical solutions at no extra cost. Second, our wavefront construction provides a simple mechanism to include a variety of optimization processes, e.g. error minimization, for a posteriori selection of optimal initial parameters. Finally, since the geometrical optics solution can be recovered by the first set of canonical solutions, it is possible to design an efficient hybrid method which switches between the geometrical optics (which does not require the post-processing step) and Gaussian beam solutions smoothly. We present numerical examples to verify the efficiency, accuracy, and the flexibility of the algorithm.

The first step of our algorithm in part 1, which is the computation of wavefronts, is an adaptation of the front tracking scheme in [34]. It is to be noted that in order to control the resolution of wavefronts, we can also adapt and include other front tracking methods, such

as the grid-based particle method [35] and the fast interface tracking method [36, 37], in the algorithm. The main contributions of this paper include the second step of the algorithm in part 1, i.e. the construction of canonical functions, and the fast post-processing technique in part 2 based on an optimal selection of the beams' initial data.

The rest of the paper is organized as follows. In Section 2 we first review the Gaussian beam models for the computation of time harmonic high frequency waves (Sections 2.1–2.4). We then present and discuss different choices of initial parameters in the computation of Gaussian beams (Section 2.5). Next, in Section 3 we describe the new wavefront method based on Gaussian beam summation and canonical functions. Numerical examples are performed in Section 4. Finally, we summarize our conclusions in Section 5.

2. Gaussian beam models

Gaussian beams are asymptotic solutions of linear wave equations. They can also be extended to some dispersive wave equations like the Schrödinger equation. Gaussian beam summation is an approximate model for linear high frequency wave propagation problems. In this approach, the initial/boundary data are decomposed into individual Gaussian beams, which are computed by a system of ODEs along their central rays. The contribution of each beam close to its central ray is approximated by Taylor expansion. The wave field is then obtained by summing over the beams. In this section, we review the governing equations for computing Gaussian beams and formulate the beam summation model.

2.1. High frequency waves and asymptotic approximations

We start with the scalar wave equation

$$v_{tt}(t,\boldsymbol{x}) - c(\boldsymbol{x})^2\,\Delta v(t,\boldsymbol{x}) = 0, \qquad (t,\boldsymbol{x}) \in \mathbb{R}_+ \times \mathbb{R}^2, \tag{1}$$

where $v = v(t,\boldsymbol{x})$ is the wave solution, t and $\boldsymbol{x} = (x,y)^\top$ are the temporal and spatial variables, respectively, and $c(\boldsymbol{x})$ is the local speed of wave propagation in the medium. We complement the wave equation (1) with highly oscillatory initial data that generate high-frequency solutions. The exact form of the data will not be important here, but a typical example is $v(0,\boldsymbol{x}) = a(\boldsymbol{x})\exp(i\,\omega\,\boldsymbol{k}\cdot\boldsymbol{x})$, where $\omega \gg 1$ is the angular frequency and $|\boldsymbol{k}| = 1$. We assume that the wavelength, which is inversely proportional to ω, is much smaller than the typical scale of the medium structure (variations in the wave speed) and the wave propagation distance (the size of the computational domain). Hence, we encounter a multiscale problem with highly oscillatory solutions. Note that with slight modifications, the techniques we describe here will also carry over to systems of wave equations, such as the Maxwell and elastodynamic equations.

We consider time-harmonic waves of type

$$v(t,\boldsymbol{x}) = u(\boldsymbol{x})\exp(i\,\omega\,t). \tag{2}$$

Inserting the ansatz (2) into the time-dependent wave equation (1), we obtain the reduced wave equation in the frequency domain, known as Helmholtz equation,

$$\Delta u(\boldsymbol{x}) + \frac{\omega^2}{c(\boldsymbol{x})^2}\,u(\boldsymbol{x}) = 0, \qquad \boldsymbol{x} \in \mathbb{R}^2. \tag{3}$$

Direct numerical simulations of (3) are very expensive, since a large number of grid points or elements is required to resolve the wave oscillations. The computational cost for fixed accuracy therefore grows at least with the rate $\mathcal{O}(\omega^2)$ in two dimensions. Consequently, in the high frequency regime, direct simulations are not feasible. To circumvent this difficulty, approximate high frequency asymptotically valid methods are often employed. They are based on constructing asymptotic Wentzel–Kramers–Brillouin (WKB) expansions of the solution in inverse powers of ω:

$$u(\boldsymbol{x}) = A(\boldsymbol{x},\omega)\, e^{i\,\omega\,\phi(\boldsymbol{x})},\ A(\boldsymbol{x},\omega) = \sum_{k=0}^{\infty} A_k(\boldsymbol{x})\,(i\,\omega)^{-k} = \sum_{k=0}^{K} A_k(\boldsymbol{x})\,(i\,\omega)^{-k} + \mathcal{O}(\omega^{-(K+1)}). \quad (4)$$

In this expansion, the phase ϕ and amplitudes A_k are independent of frequency and vary on a much coarser scale than the full wave solution. They can therefore be computed at a computational cost independent of the frequency.

2.2. Geometrical optics

One popular asymptotic method is geometrical optics (GO) [1, 38]. In its standard form, it only considers the leading term of the series ($K = 0$), called the the geometrical optics term. It introduces an error of order $\mathcal{O}(\omega^{-1})$. The phase ϕ and amplitude A_0 are real-valued functions and satisfy the eikonal and transport PDEs respectively,

$$|\nabla\phi|^2 = 1/c(\boldsymbol{x})^2, \qquad 2\,\nabla\phi\cdot\nabla A_0 + A_0\,\Delta\phi = 0. \quad (5)$$

GO can also be formulated in terms of ODEs. The eikonal equation is a nonlinear Hamilton-Jacobi equation with Hamiltonian $H(\boldsymbol{x},\boldsymbol{p}) = c(\boldsymbol{x})\,|\boldsymbol{p}| \equiv 1$, where $\boldsymbol{p} = \nabla\phi$ is called the slowness vector. We let $(\boldsymbol{x}(t),\boldsymbol{p}(t))$ be a bi-characteristic related to this Hamiltonian, satisfying the so called ray equations,

$$\frac{d\boldsymbol{x}}{dt} = \nabla_{\boldsymbol{p}} H = c^2\,\boldsymbol{p}, \qquad \frac{d\boldsymbol{p}}{dt} = -\nabla_{\boldsymbol{x}} H = -\frac{\nabla c}{c}. \quad (6)$$

The parameter t corresponds to the waves' travel time in the sense that $\phi(\boldsymbol{x}(t)) = \phi(\boldsymbol{x}(0)) + t$. There are also ODEs for the amplitude [1]. The main drawback of geometrical optics is that the approximation breaks down at caustics, where rays concentrate and the predicted amplitude is unbounded [39].

2.3. Gaussian beams

Gaussian beams constitute another high frequency asymptotic model which is closely related to GO and yet is valid at caustics. The solution is assumed to be of the same form (4), but there are two important differences. First, while a GO solution is globally defined for all rays, a Gaussian beam is a localized solution that concentrates near a single ray of GO, known as the beam's central ray and denoted by $\boldsymbol{x}^*(t)$. Secondly, while in GO the phase is real-valued, in the Gaussian beam construction it is real-valued only on the central ray of the beam. Away from the central ray, it is complex-valued with positive imaginary part.

The solution will then be exponentially decreasing away from the central ray, maintaining its Gaussian shape. If we only take the first term in (4), we write the Gaussian beam solution as

$$u_{\mathrm{GB}}(\boldsymbol{x}) = A_0(\boldsymbol{x})\, e^{i\,\omega\,\phi(\boldsymbol{x})}. \tag{7}$$

Because of the localization, one can approximate the complex-valued ϕ and A_0 close to the central ray (i.e. when $\boldsymbol{x}$ is close to $\boldsymbol{x}^*(t)$) by Taylor expansions around the ray. For instance, the simplest "first order" Gaussian beam is constructed by

$$\phi(\boldsymbol{x}) \approx \phi(\boldsymbol{x}^*) + (\boldsymbol{x}-\boldsymbol{x}^*)\cdot\nabla\phi(\boldsymbol{x}^*) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^*)^\top D^2\phi(\boldsymbol{x}^*)\,(\boldsymbol{x}-\boldsymbol{x}^*), \tag{8}$$

$$A_0(\boldsymbol{x}) \approx A_0(\boldsymbol{x}^*), \tag{9}$$

i.e. ϕ and A_0 are approximated to 2nd and 0th orders with respect to $\boldsymbol{x}-\boldsymbol{x}^*$, respectively. For a given $\boldsymbol{x}$ there is some freedom in the choice of point $\boldsymbol{x}^*$ on the central ray, around which the Taylor expansion should be made. Typically, $\boldsymbol{x}^*$ is taken as the point on the ray that is closest to $\boldsymbol{x}$, but other choices are also possible, e.g. the closest point with the same x- or y-coordinate.

The beam's central ray $\boldsymbol{x}^*(t) = (x^*, y^*)^\top$ is given by the *ray tracing* equations

$$\begin{aligned}
\frac{dx^*}{dt} &= c(\boldsymbol{x}^*)\,\cos\theta^*, & x^*(0) &= x_0^*,\\
\frac{dy^*}{dt} &= c(\boldsymbol{x}^*)\,\sin\theta^*, & y^*(0) &= y_0^*,\\
\frac{d\theta^*}{dt} &= \partial_x c(\boldsymbol{x}^*)\,\sin\theta^* - \partial_y c(\boldsymbol{x}^*)\,\cos\theta^*, & \theta^*(0) &= \theta_0^*,
\end{aligned} \tag{10}$$

which are obtained from (6) by setting $\boldsymbol{p} = \nabla\phi = (\cos\theta, \sin\theta)^\top / c(\boldsymbol{x})$, thanks to the eikonal equation in (5), with θ being the angle between the tangent of the ray and the positive x-axis. In fact, the slowness vector $\boldsymbol{p}$ represents the direction of the central ray. The initial conditions in (10) are given by the initial location $\boldsymbol{x}_0^* = (x_0^*, y_0^*)$ and initial angle θ_0^* of the central ray. After computing the central ray from (10), the Taylor coefficients $\phi(\boldsymbol{x}^*)$, $\nabla\phi(\boldsymbol{x}^*)$, $D^2\phi(\boldsymbol{x}^*)$, and $A_0(\boldsymbol{x}^*)$ in (8) and (9) are then computed only on the central ray using the transport equation in (5) [7]:

$$\phi(\boldsymbol{x}^*) = \phi(\boldsymbol{x}^*(0)) + t, \qquad \nabla\phi(\boldsymbol{x}^*) = (\cos\theta^*, \sin\theta^*)^\top / c(\boldsymbol{x}^*), \qquad D^2\phi(\boldsymbol{x}^*) = TMT^{-1},$$
$$A(\boldsymbol{x}^*) = (c(\boldsymbol{x}^*)/Q)^{1/2},$$

where

$$T = \begin{pmatrix} \sin\theta^* & \cos\theta^* \\ -\cos\theta^* & \sin\theta^* \end{pmatrix}, \quad M = \begin{pmatrix} P/Q & -c_1/c(\boldsymbol{x}^*)^2 \\ -c_1/c(\boldsymbol{x}^*)^2 & -c_2/c(\boldsymbol{x}^*)^2 \end{pmatrix}, \quad \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = T^{-1}\nabla c(\boldsymbol{x}^*).$$

The complex-valued scalar functions P and Q are given by the *dynamic ray tracing* equations

$$\begin{aligned}
\frac{dQ}{dt} &= c(\boldsymbol{x}^*)^2\, P, & Q(0) &= Q_0,\\
\frac{dP}{dt} &= g(\boldsymbol{x}^*, \theta^*)\, Q, & P(0) &= P_0,
\end{aligned} \tag{11}$$

where

$$g(\boldsymbol{x}^*, \theta^*) = -\frac{\partial_{xx} c(\boldsymbol{x}^*)\, \sin^2\theta^* - 2\partial_{xy} c(\boldsymbol{x}^*)\, \sin\theta^*\, \cos\theta^* + \partial_{yy} c(\boldsymbol{x}^*)\, \cos^2\theta^*}{c(\boldsymbol{x}^*)}. \tag{12}$$

The quantities P and Q determine the wavefront curvature and the beam width.

First order Gaussian beams have an asymptotic error of size $\mathcal{O}(1/\sqrt{\omega})$. Beams that are higher order accurate in ω can be constructed by taking more terms both in the WKB expansion (4) and in the Taylor expansions of ϕ and A.

2.4. Gaussian beam summation

Since (3) is a linear equation, it is a natural extension to consider a superposition of Gaussian beams to represent more general high frequency solutions that are not necessarily localized around a single ray. We therefore consider wave fields generated by a wave source $u_0(\boldsymbol{x})$ which concentrates on a curve $\boldsymbol{x}_0(s)$ in $\mathbb{R}^2$ parameterized by the arc length parameter s. Mathematically, this corresponds to finding the outgoing solution of

$$\Delta u(\boldsymbol{x}) + \frac{\omega^2}{c(\boldsymbol{x})^2}\, u(\boldsymbol{x}) = 2\, i\, \omega\, u_0(\boldsymbol{x})\, \delta_{\boldsymbol{x}_0}(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^2. \tag{13}$$

where the Dirac delta function $\delta_{\boldsymbol{x}_0}(\boldsymbol{x})$ is supported on $\boldsymbol{x}_0(s)$, and $u_0(\boldsymbol{x})$ is a real-valued smooth, compactly supported and non-oscillatory function. We note that by setting u_0 to a real-valued function, we enforce waves to propagate in the normal direction to $\boldsymbol{x}_0(s)$. By setting u_0 to an oscillatory complex-valued function, we can include more general wave sources which may not necessarily propagate in orthogonal directions to $\boldsymbol{x}_0(s)$. Here, we only consider real-valued u_0 for simplicity. Under suitable conditions on $c(\boldsymbol{x})$, problem (13) is well-posed and has a well-defined high-frequency limit where $u = u_0$ on $\boldsymbol{x}_0(s)$; see [18] for more details. The curve $\boldsymbol{x}_0(s)$ can be interpreted as the initial wavefront of the high-frequency solution.

We introduce the notation $A(\boldsymbol{x}, s)$ and $\phi(\boldsymbol{x}, s)$ for the amplitude and phase of a beam with the initial position $\boldsymbol{x}_0(s)$ of its central ray. The central rays are denoted by $\boldsymbol{x}^*(t, s) = (x^*(t, s), y^*(t, s))$, with angle $\theta^*(t, s)$. They satisfy (10) where the initial position is $\boldsymbol{x}^*(0, s) = \boldsymbol{x}_0(s)$, and the initial angle $\theta^*(0, s)$ is chosen such that the central ray initially points normal to $\boldsymbol{x}_0(s)$. Since u_0 is real, we also let $\phi(\boldsymbol{x}_0(s), s) = 0$ for all s.

We decompose the wave field generated by u_0 on $\boldsymbol{x}_0(s)$ into $N_{\mathrm{GB}} \in \mathbb{N}$ beams by discretizing s, setting $s_m = m\, h$ with $m = 1, \ldots, N_{\mathrm{GB}}$, where h is a small arc length representing the beam spacing on the curve $\boldsymbol{x}_0(s)$, see e.g. [7]. The initial locations of the beams' central rays are $\{\boldsymbol{x}_0(s_m)\}_{m=1}^{N_{\mathrm{GB}}}$. The initial angles $\theta_0(s_m)$ of the central rays, pointing normal to $\boldsymbol{x}_0(s)$, are also obtained for all discretization points. With these initial conditions, the central rays of N_{GB} beams are then computed by solving N_{GB} ray tracing systems (10). In order to compute the individual Gaussian beams, we also need to solve N_{GB} dynamic ray tracing systems (11) with corresponding initial conditions. Similarly to above, the dynamic ray tracing quantities in (11) are denoted $P(t, s)$ and $Q(t, s)$, with the corresponding initial data $P_0(s)$, $Q_0(s)$. After computing these quantities along N_{GB} central rays, the contribution of the beams concentrated close to their central rays are determined by the approximations (8)

and (9) entered in (7). Eventually, the wave field at a fixed receiver point $\boldsymbol{x}_R$ is calculated by summing over the beams

$$u(\boldsymbol{x}_R) = \sum_{m=1}^{N_{GB}} \psi(s_m)\, A(\boldsymbol{x}_R, s_m)\, e^{i\omega\phi(\boldsymbol{x}_R, s_m)}. \tag{14}$$

In order to obtain the final approximation by (14), we must select the weights $\psi(s_m)$ and the initial data $P_0(s_m)$ and $Q(s_m)$ in (11) such that the weighted sum u in (14) well approximates u_0 on $\boldsymbol{x}_0(s)$, i.e. $u(\boldsymbol{x}_0(s)) \approx u_0(\boldsymbol{x}_0(s))$. In particular, the choice of the initial data for dynamic ray tracing systems is not straightforward and will be discussed further in Section 2.5.

In beam summation, the approximation error depends both on ω and on the spacing h of the beams. For first order beams, the convergence rate, with perfect approximation of data u_0, was shown in [18] to be at least $\mathcal{O}(1/\sqrt{\omega})$ in the limit $h \to 0$. However, numerical evidence and theory for simplified settings in [17] suggest the rate $\mathcal{O}(1/\omega + \exp(-C/h\sqrt{\omega}))$. Hence, the convergence rate in ω is actually faster than for the individual beams. Moreover, the spacing h must be taken small at high frequencies, $h \sim 1/\sqrt{\omega}$, but the convergence in $h\sqrt{\omega}$ is exponentially fast.

2.5. Initial data for dynamic ray tracing equations

There are many choices of initial conditions P_0, Q_0 and weights ψ that lead to an accurate solution, which converges to the exact solution as $h \to 0$ and $\omega \to \infty$, with the same rate as when the data u_0 is perfectly approximated. However, the choices are far from equivalent, as the corresponding pre factors in the error estimates can vary considerably, leading to large differences in accuracy. Moreover, even if two different choices give almost the same accuracy on the initial curve $\boldsymbol{x}_0(s)$, they can generate a significantly different solution quality away from $\boldsymbol{x}_0(s)$. A main difficulty in the Gaussian beam method is therefore to select the data P_0, Q_0 and ψ such that the error in the Gaussian beam solution is minimized [6, 40, 17, 33, 41]. The optimal choice also depends on where the solution is sought.

Not all values of P_0 and Q_0 give meaningful wave fields. It can be shown that to be admissible, the parameters must satisfy

$$Q_0 \neq 0, \qquad \Im(P_0/Q_0) > 0. \tag{15}$$

Then $Q(t) \neq 0$ and $\Im(P(t)/Q(t)) > 0$ for all $t > 0$ along the central ray [3]. The former guarantees the *regularity* of the Gaussian beam (with finite amplitudes at caustics), and the latter guarantees the *non-degeneracy* of the beam (concentration of the solution close to the ray).

The primary free variable is Q_0, and we will only consider $P_0 = i$ in what follows. Together, P_0 and Q_0 determine ψ_j. The admissibility condition (15) then reduces to simply

$$\Re Q_0 > 0. \tag{16}$$

We start with the typical case of a plane wave source.

Example 1. *Decomposition of a plane wave into Gaussian beams.* Consider the wave field generated by a plane wave at $\boldsymbol{x}_0(s) = (0, s)$ propagating into the domain $x > 0$ orthogonally, i.e. $\theta^*(0, s) = 0$, with $c \equiv 1$ in a neighborhood of $\boldsymbol{x}_0$. This means that $u_0 \equiv 1$ and we should find Q_0, P_0 and ψ such that $u(0, y) \approx 1$. We will show that, for each h and ω, there is a family of possible ways to accurately approximate u_0, parameterized by Q_0.

For a fixed s, we have on the initial curve

$$\phi(\boldsymbol{x}_0(s), s) = 0, \qquad \nabla\phi(\boldsymbol{x}_0(s), s) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad D^2\phi(\boldsymbol{x}_0(s), s) = \begin{pmatrix} 0 & 0 \\ 0 & P_0(s)/Q_0(s) \end{pmatrix}.$$

Using the closest point for Taylor expansion, $(0, y) - \boldsymbol{x}^* = (0, y) - \boldsymbol{x}_0(s) = (0, y - s)$, we get from (8) and (9),

$$\phi(0, y, s) = \frac{1}{2}(y - s)^2 \frac{P_0(s)}{Q_0(s)}, \qquad A(\boldsymbol{x}_0(s), s) = \frac{1}{\sqrt{Q_0(s)}},$$

and the wave field (14) at the initial curve is

$$u(0, y) = \sum_m \psi(s_m) \frac{1}{\sqrt{Q_0(s_m)}} e^{\frac{i}{2}\omega(y - s_m)^2 \frac{P_0(s_m)}{Q_0(s_m)}}. \tag{17}$$

We now note that the constant function can be well approximated by a weighted sum of gaussians [10]. Indeed, for all $y \in \mathbb{R}$,

$$1 = \sum_m \frac{1}{\sqrt{\pi}} \frac{h}{\mathsf{w}_0} e^{-(y - s_m)^2/\mathsf{w}_0^2} + \mathcal{O}\left(e^{-(\mathsf{w}_0/h)^2}\right), \quad s_m = m\, h, \tag{18}$$

with h and w_0 representing the spacing of the gaussians and their half-widths; see Figure 1. To properly choose the initial data in (17), we identify the two expressions and use $P_0 = i$.

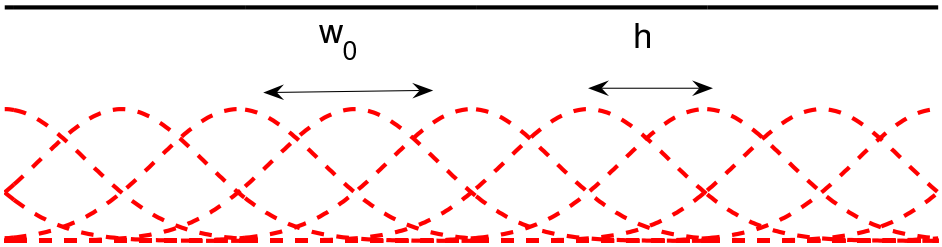


Figure 1: The sum of several translated gaussians is almost constant. A plane wave can therefore be decomposed approximately to a sum of parallel Gaussian beams.

The Gaussian beam solution (17) produces a plane wave for any $Q_0(s_j) = Q_0$ if

$$\mathsf{w}_0 = \left(\frac{2Q_0}{\omega}\right)^{1/2}, \qquad \psi(s_m) = h\left(\frac{\omega}{2\pi}\right)^{1/2}.$$

There are some restrictions on the possible choices of Q_0. For accuracy, i.e. to have a small error term in (18), one must take the parameters w_0 and h such that $\mathsf{w}_0 > h$, i.e. $Q_0 > \omega h^2/2$. This means that w_0 cannot be too small since, for computational efficiency, $h \propto 1/N_{GB}$ should not be too small. Similarly, w_0 should not be too big, since the evaluation of the beams become expensive when they are wide. Moreover, (18) is in fact valid also for complex-valued w_0 with the error term replaced by $\mathcal{O}(\exp(-\Re \mathsf{w}_0^2/h^2))$. We can therefore also allow complex-valued Q_0 as long as (15) is satisfied and $\Re \mathsf{w}_0^2 > h^2$, which simply means that $\Re Q_0 > \omega h^2/2$. We note finally that if we use different Q_0 at different points, $Q_0 = Q_0(s_j)$, the expansion (18) is valid with an additional approximation error of size $\mathcal{O}(\mathsf{w}_0^2)$.

The case of the plane wave is generic in the sense that Q_0 is in general a free parameter. The question to be answered is how to select Q_0 optimally. It has been proposed that the optimal choice should produce Gaussian beams of minimum width at the receiver point, see e.g. [6, 40]. One should then find Q_0 which minimizes the half-width, normal to the central ray, of the Gaussian beam, which is given by

$$\mathsf{w}(t,s) = \left(\frac{1}{2}\omega\, \Im \partial_n^2 \phi(\boldsymbol{x}^*(t,s),s)\right)^{-1/2} = \left(\frac{1}{2}\omega\, \Im (P(t,s)/Q(t,s))\right)^{-1/2}. \tag{19}$$

The main motivation for this choice is that for wide beams, the Taylor expansion error should be large. Moreover, from the computational point of view, it is more convenient to work with beams which are as narrow as possible, because in the case of variable speed of propagation, where the central rays can bend, at some distance from the rays the closest point on the ray $\boldsymbol{x}^*$ may become non-unique; the phase then becomes non-smooth and the Gaussian beam approximation breaks down. However, it was shown in [17, 33] that optimally narrow beams will not necessarily give the minimum error. In particular, White et al. [33] study different choices of initial data. For particular types of problems, they obtain initial data which give a more accurate approximation compared to the data obtained by the minimization of the beam width. In general, the optimal choice of Q_0 should minimize the error, not necessarily the width. How to find it is still an open question.

As noted above, the optimal Q_0 depends on the location of the receiver point, i.e. where the beam summation (14) is evaluated. The optimal value is different for different beams and different points along a beam's central ray, meaning it depends both on t and s. To find optimal solutions at different locations, we may therefore need to solve (11) for many different initial conditions P_0 and Q_0. This can be computationally very expensive. However, we can take advantage of the linearity of (11) and make the following important observation. We specify two real-valued functions (Q_I, P_I) and (Q_{II}, P_{II}), called *canonical solutions*, which satisfy (11) with two different sets of initial data:

$$(Q_I, P_I)(0) = (1,0), \quad (Q_{II}, P_{II})(0) = (0,1).$$

Consequently, the two canonical solutions satisfy the following two ODE systems

$$\begin{aligned} \frac{dQ_I}{dt} &= c(\boldsymbol{x}^*)^2\, P_I, \qquad Q_I(0) = 1, & \qquad \frac{dQ_{II}}{dt} &= c(\boldsymbol{x}^*)^2\, P_{II}, \qquad Q_{II}(0) = 0, \\ \frac{dP_I}{dt} &= g(\boldsymbol{x}^*, \theta^*)\, Q_I, \quad P_I(0) = 0, & \qquad \frac{dP_{II}}{dt} &= g(\boldsymbol{x}^*, \theta^*)\, Q_{II}, \quad P_{II}(0) = 1, \end{aligned} \tag{20}$$

where g is given by (12). Then, the complex-valued solution (Q, P) of (11) with the initial data (Q_0, P_0) is given by [6, 33]:

$$Q = Q_0\, Q_I + P_0\, Q_{II}, \qquad P = Q_0\, P_I + P_0\, P_{II}. \tag{21}$$

Hence, from two canonical solutions, beams with all possible initial data can be computed by taking linear combinations at no extra cost. Note that the canonical solutions also depend on the parameter s via $\boldsymbol{x}^*(t, s)$, and when needed we write $Q_I(t, s)$, etc. to indicate this.

We also note that, in particular, the geometrical optics solution can be obtained from the first canonical solution Q_I,

$$\phi_{\mathrm{GO}}(\boldsymbol{x}^*(t)) = \phi(\boldsymbol{x}^*(0)) + t, \qquad A_{\mathrm{GO}}(\boldsymbol{x}^*(t)) = A(\boldsymbol{x}^*(0)) \left(\frac{1}{Q_I(t)} \frac{c(\boldsymbol{x}^*(t))}{c(\boldsymbol{x}^*(0))} \right)^{1/2}, \tag{22}$$

which corresponds to an infinitely wide beam.

There are two major advantages of using canonical solutions:

1. *Optimization.* Canonical solutions provide an efficient mechanism for performing and including the optimization of initial parameters. Currently, the process of computing the optimal initial data that minimize the error is still an open question and needs further investigations. In this paper, we instead propose a simple strategy based on the minimization of beam widths to find suitable initial parameters (see Example 2 below and the numerical tests in Section 4). The key point here is that any optimization process, based on the minimization of either beam width or error, can simply be included in the algorithm, thanks to (20) and (21).

2. *Hybridization.* Since the geometrical optics solution can be recovered by the first set of canonical solutions by (22), it is possible to design a hybrid method which switches between the geometrical optics and Gaussian beam solutions smoothly. Such a hybrid technique would substantially reduce the computational complexity, because in places where there are no caustics, the computation of Gaussian beams (the post-processing part in Algorithm 1 below) is not needed.

Example 2. *Optimal selection of initial parameter Q_0.* In order to clarify the first advantage, we present a simple optimization procedure and show how to obtain "optimal" initial conditions for (11). Our optimization is based on minimizing the beam width at (t, s) with $P_0 = i$. In general, the best choice of Q_0 will vary with both t and s. Consider a possibly complex-valued $Q_0 = q_r + q_i\, i$ with $q_r > 0$, so that condition (16) is satisfied. To find $Q_0^{\min}(t, s)$ corresponding to the narrowest possible beam, we write by (19) and (21),

$$Q_0^{\min}(t, s) = \operatorname*{argmin}_{Q_0} \mathsf{w}(t, s) = \operatorname*{argmax}_{Q_0} \Im \frac{P(t, s)}{Q(t, s)} = \operatorname*{argmax}_{q_r, q_i} \frac{q_r (Q_I\, P_{II} - P_I\, Q_{II})}{q_r^2\, Q_I^2 + (q_i\, Q_I + Q_{II})^2}. \tag{23}$$

We consider two cases. First, we let Q_0 be real-valued and set $q_i = 0$. We then obtain q_r by solving the minimization problem (23). The optimal Q_0 is

$$Q_0^{\min}(t, s) = \left| \frac{Q_{II}(t, s)}{Q_I(t, s)} \right|, \tag{24}$$

which gives the minimal half-width at (s,t)

$$\mathsf{w}^{\min} = \left(\frac{4|Q_{II}Q_I|}{\omega(Q_I\,P_{II} - P_I\,Q_{II})} \right)^{1/2}.$$

Next, we allow Q_0 to be complex-valued. In this case, we can make the beam width arbitrarily small. Indeed, for any $q_r > 0$ and $\alpha \in \mathbb{R}$, we can select

$$Q_0 = q_r - (1 - \alpha\sqrt{q_r})\frac{Q_{II}(t,s)}{Q_I(t,s)}\,i. \tag{25}$$

and obtain the half-width

$$\mathsf{w}(t,s) = \left(2q_r \frac{Q_I^2 + \alpha^2 Q_I^2}{\omega(Q_I\,P_{II} - P_I\,Q_{II})} \right)^{1/2},$$

which can be made as small as we like by taking q_r small. As $q_r \to 0$, we get an infinitely *wide* initial beam, however. Still, a complex-valued Q_0 gives more control over the beam character. See Section 4 where we easily improve the accuracy of the approximate solutions by choosing a complex-valued Q_0, which generates narrower, more accurate, beams compared to a real-valued Q_0.

3. Wavefront-based Gaussian beam method

The usual way to compute high frequency wave fields by Gaussian beam summation is based on standard ray tracing, where the central rays of the beams are traced individually by solving the ODE systems (10) and (11). The main problem with ray tracing is that it may produce diverging rays that fail to cover the computational domain. In this case, one needs to increase the number of rays, which in turn increases the computational cost.

In standard geometrical optics, the problem of diverging rays can be overcome by instead using so-called wavefront methods [30, 31, 34, 42]. They are related to ray tracing, but instead of tracing a sequence of individual rays, a wavefront is evolved in physical or phase space according to the ODE formulations. In physical space, a wavefront at a travel-time $t \geq 0$ is a curve $\{\boldsymbol{x}(t,s) \mid \phi(\boldsymbol{x}(t,s),s) - \phi(\boldsymbol{x}_0(s),s) - t = 0\}$, i.e. an iso-phase line. Wavefront methods provide a simple mechanism for controlling the resolution and accuracy of the numerical approximation.

Using wavefront methods with Gaussian beams is not as straightforward, since the beam method strongly depends on the distribution and width of the beams at the initial front. To clarify this difficulty, let us consider a beam's central ray $\boldsymbol{x}^*(t)$ starting at $\boldsymbol{x}^*(0)$. As discussed in Section 2.5, it is not wise to choose the same initial conditions (Q_0, P_0) for all travel-times $t > 0$ along the ray. A fixed initial data set generates different beam widths at different points on the central ray. Consequently, at some travel-times along the ray, we may get very wide beams, which generate very large errors in Taylor approximations. A naive way to get around this problem is to compute $(Q(t), P(t))$ at each travel-time and for each individual beam use a particular set of initial data that generates the narrowest beam or the smallest

error at that particular time and for that particular beam. This, however, requires solving many ODE systems (11) with many different initial data, which is prohibitively expensive. We therefore need an efficient strategy to allow the computation of $(Q(t), P(t))$ at different travel-times $t > 0$ with different initial conditions $(Q(0), P(0))$.

We introduce a Lagrangian wavefront-based Gaussian beam method, in which a wavefront is evolved in phase space $(\boldsymbol{x}, \theta)$ by solving the ODE systems (10) and (11). The proposed method consists of three major steps. First, in order to overcome the problem of diverging rays in the ray tracing method, we use an automatic refinement criterion to keep the fronts uniformly sampled, similar to [34]. This will result in an *adaptive front tracking* scheme. Secondly, in order to account for the problem with initial conditions explained above, we propose an efficient strategy based on the *evolution of canonical solutions* (Q_I, P_I) and (Q_{II}, P_{II}). Finally, we perform a post-processing step based on the Gaussian beam summation technique to compute the wave field at any desired receiver point. We will now explain the three steps of the algorithm in more detail.

3.1. Adaptive front tracking

We let the initial phase space wavefront be $(\boldsymbol{x}_0(s), \theta_0(s))$ parameterized by s, and denote as before the exact phase space wavefront at travel-time t by $(\boldsymbol{x}(t, s), \theta(t, s))$. Next, we introduce the numerical approximations

$$\boldsymbol{x}_j^n \approx \boldsymbol{x}(n\,\Delta t, j\,\Delta s), \qquad \theta_j^n \approx \theta(n\,\Delta t, j\,\Delta s),$$

where (j, n) represents a marker (grid point) on a front at $t = n\,\Delta t$. Note that the time and space steps, Δt and Δs, do not need to resolve the high frequency wave lengths, and in general they are much bigger than $1/\omega$. We initialize $N_0 \in \mathbb{N}$ markers on the initial front at $t = 0$ as $(\boldsymbol{x}_j^0, \theta_j^0) = (\boldsymbol{x}_0(j\,\Delta s), \theta_0(j\,\Delta s))$, with $j = 1, 2, \ldots, N_0$. Each marker is then updated by a standard ODE-solver, such as the Runge-Kutta method, applied to the ray tracing system (10). See Figure 2 (left).

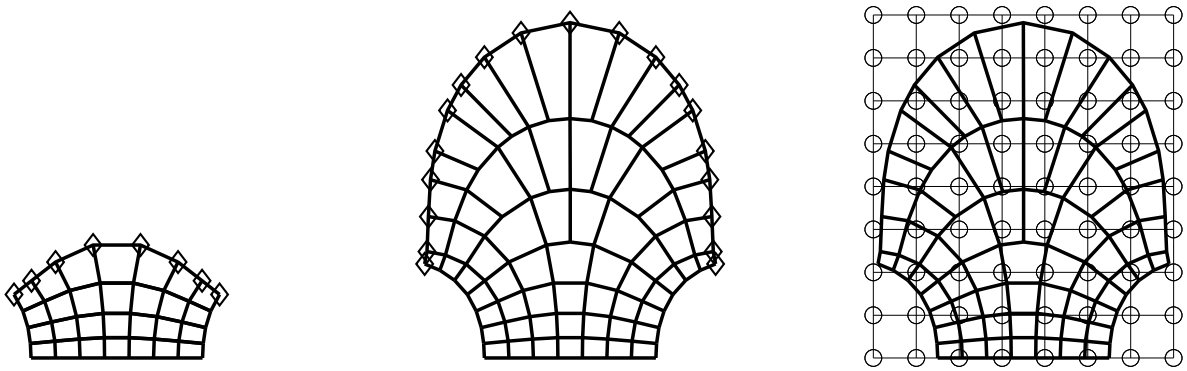


Figure 2: Wavefront construction. Markers ($\diamond$) on the wavefront are propagated along the ordinary rays (left). New markers are inserted to keep a uniform sampling of the front (middle). The information carried by markers are interpolated onto a regular grid (right).

When the resolution of the wavefront deteriorates, new markers are inserted and computed by interpolation from the old markers. We add a new marker $(j+1/2, n)$ between markers (j, n) and $(j+1, n)$ if

$$|\boldsymbol{x}_{j+1}^n - \boldsymbol{x}_j^n| \geq \delta_{\boldsymbol{x}} \qquad or \qquad |\theta_{j+1}^n - \theta_j^n| \geq \delta_\theta,$$

for some tolerances $\delta_{\boldsymbol{x}}$ and δ_θ. See Figure 2 (middle). Eventually, we will have $N_{\rm w} \geq N_0$ markers on the last wavefront.

In order to illustrate the importance of the above adaptive refinement strategy, we perform a simple numerical test carried out both with and without adaptive refinement. A plane wave propagates into the domain from the left boundary. Figure 3 shows the central rays (top row) and θ versus y along the front $t = 3.5$ (bottom row) obtained by the wavefront method. As it can be seen, in the case of no refinement (left column), the solution is poorly resolved in places where the rays diverge. However, the solution is well resolved if refinement is performed (right column).

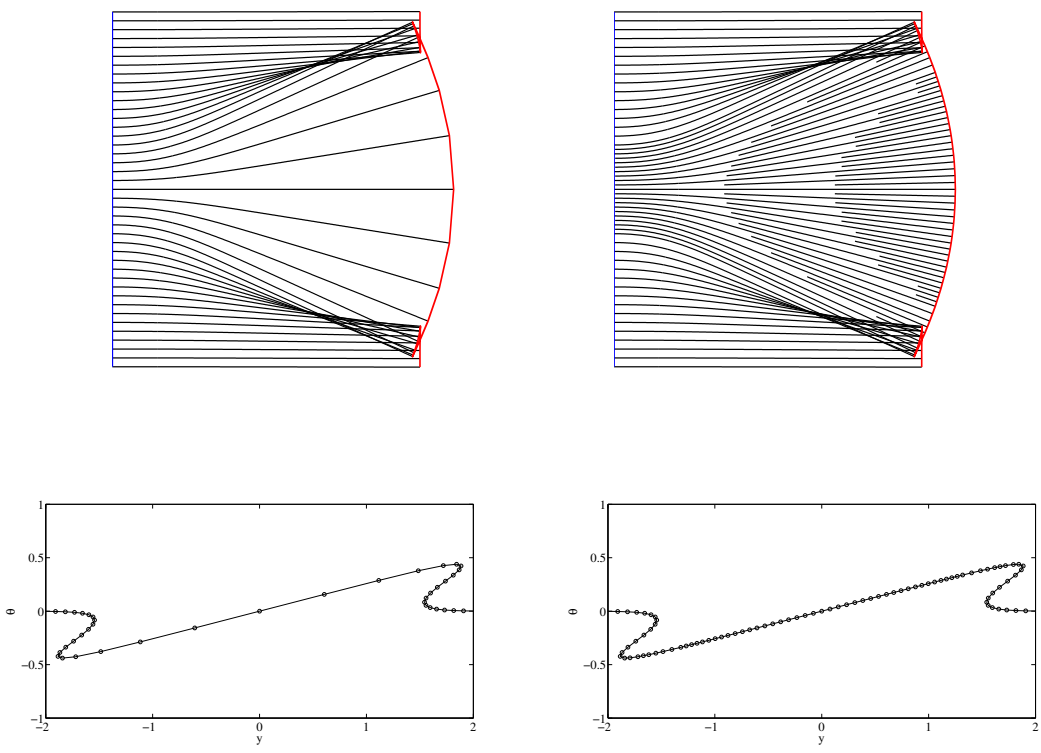


Figure 3: A plane wave propagates into the domain from the left boundary. Top figures show the beam central rays computed without refinement (left) and with refinement (right). Thick curves are the fronts at $t = 3.5$. Bottom figures show the corresponding θ functions versus y along the front $t = 3.5$. In places where the rays diverge, the non-refined solution is poorly resolved, while the refined solution is uniformly resolved.

Note that inserting new markers on the fronts in the wavefront method is analogous to inserting new rays in the ray tracing method. However, here, the rays are inserted only in places where the resolution deteriorates. These rays are traced only after this point in time, and there is no need to compute them from the source, as is done in the ray tracing method. Therefore, the wavefront method is computationally faster than the ray tracing method, while keeping the same accuracy.

3.2. Evolution of canonical solutions

In parallel with computing $(\boldsymbol{x}_j^n, \theta_j^n)$, we also compute the corresponding real-valued canonical functions $(Q_{I_j}^n, P_{I_j}^n)$ and $(Q_{II_j}^n, P_{II_j}^n)$ by solving the dynamic ray tracing systems (20) with $(\boldsymbol{x}_j^n, \theta_j^n)$ and fixed initial conditions. When new marker points are added by interpolation, new values of $(Q_{I_j}^n, P_{I_j}^n)$ and $(Q_{II_j}^n, P_{II_j}^n)$ are also interpolated and added. We note that via (21) we can recreate beams with *any* initial data P_0 and Q_0 from these two canonical solutions. We save $V_j^n := (x_j^n, y_j^n, \theta_j^n, Q_{I_j}^n, P_{I_j}^n, Q_{II_j}^n, P_{II_j}^n)^\top \in \mathbb{R}^7$ for each grid point (j, n).

3.3. Post-processing by Gaussian beam summation

Now assume we want to compute the wave field at a set of receiver points $\{\boldsymbol{x}_R\}$ on the front $\boldsymbol{x}(t, s)$ at $t = t^* = n^* \Delta t$. We first select the initial spacing h of the beams, the beam parameters (Q_0, P_0), and the weights ψ such that the wave source u_0 is well approximated on the initial front and such that the beam widths on the front at $t = t^*$ are small (see Sections 2.4 and 2.5). The initial spacing h determines the number of beams $N_{GB} \propto 1/h$ and gives a uniform discretization of the initial front with the grid points $\{s_m\}$, where $s_m = m\, h$ and $m = 1, \ldots, N_{GB}$. For instance, if the source is given on the vertical line $\{\boldsymbol{x} \,|\, x = 0, y \in [0, 1]\}$, we get $N_{GB} = 1 + 1/h$ and obtain $\boldsymbol{x}_0(s_m) = (0, s_m - h)$. We note that $h \sim 1/\sqrt{\omega} \ll \Delta s$ so the number of such grid points N_{GB} is in general much higher than the number of marker points N_w on the wave front at $t = t^*$.

Each grid point on the initial front represents the initial point of a beam's central ray. To find the corresponding values of $\boldsymbol{x}, \theta, P, Q$ at $t = t^*$, we do not need to recompute the rays. Instead, we find approximations of $\boldsymbol{x}(t^*, s_m), \theta(t^*, s_m)$ and the canonical solutions $Q_I(t^*, s_m),\ P_I(t^*, s_m),\ Q_{II}(t^*, s_m),\ P_{II}(t^*, s_m)$, with $m = 1, \ldots, N_{GB}$, by interpolating the already computed values $V_j^{n^*}$ with $j = 1, \ldots, N_w$. The complex-valued numbers $Q(t^*, s_m)$ and $P(t^*, s_m)$ are then obtained from (21). Finally, the total wave field at the receiver points $\{\boldsymbol{x}_R\}$ is calculated by (14).

As an alternative way, if we need the wave field on a regular grid, we can first interpolate V_j^n values down on a regular Cartesian grid. See Figure 2 (right). We then use the same procedure as above, but instead of a wavefront, we consider a line passing the receiver point.

Remark 1. *We emphasize that a main advantage of the proposed algorithm based on the canonical solutions (Q_I, P_I) and (Q_{II}, P_{II}) is that at different receivers, we can use different initial data (Q_0, P_0) to evaluate the solution at no extra cost. Therefore, optimization, based on the minimization of either the beam width or the error, is possible, and we can efficiently approximate the field at different receivers using optimally chosen beams. Moreover, since the geometrical optics solutions can be obtained by Q_I, it is practically possible to construct a hybrid algorithm and use Gaussian beam solutions only around caustics.*

To summarize the above steps, we can divide the wavefront-based Gaussian beam algorithm, outlined in Algorithm 1, into two parts: (I) the computation of wave fronts and canonical solutions; and (II) the computation of Gaussian beams and the wave field. The cost of computing the wavefront and the V_j^n quantities is independent of ω. Since the beams are localized, it is possible to discard most of them in the sum (14); only a few of them contribute to the field at each receiver point. The cost to compute the field is therefore typically $\mathcal{O}(1)$ per receiver point.

Algorithm 1 Wavefront-based Gaussian beam method

Part I. *Calculate the front $t = T$ and the canonical solutions.*

0. Given a wave source u_0 on the initial front $\boldsymbol{x}_0(s)$ with initial propagation angle $\theta_0(s)$.
1. Discretize the initial front into N_0 grid points and obtain N_0 initial data for (10).
2. Evolve the front until $t = T$ by solving N_0 ODE systems (10) with adaptive refinement.
3. In parallel with **2**, solve N_0 ODE systems (20) and compute canonical solutions.
4. Collect the ODE solutions $\{(\boldsymbol{x}_j, \theta_j, Q_{Ij}, P_{Ij}, Q_{IIj}, P_{IIj})\}_{j=1}^{N_{\rm w}}$ at $t = T$ in $V_{N_{\rm w}}$.

Part II. *Post-processing.*

0. Given a set of receiver points $\{\boldsymbol{x}_{\rm R}\}$ on the front $t = T$.
1. Choose initial spacing h and the number $N_{\rm GB}$ of beams, see Example 1.
2. Interpolate $V_{N_{\rm w}}$ for $N_{\rm w}$ beams to find $V_{N_{\rm GB}}$ for $N_{\rm GB}$ beams.
3. Find (Q, P) for $N_{\rm GB}$ beams, based on proper choices of (Q_0, P_0), see Example 2.
4. Sum up the contribution of $N_{\rm GB}$ beams to calculate the wave field at the receivers.

4. Numerical examples

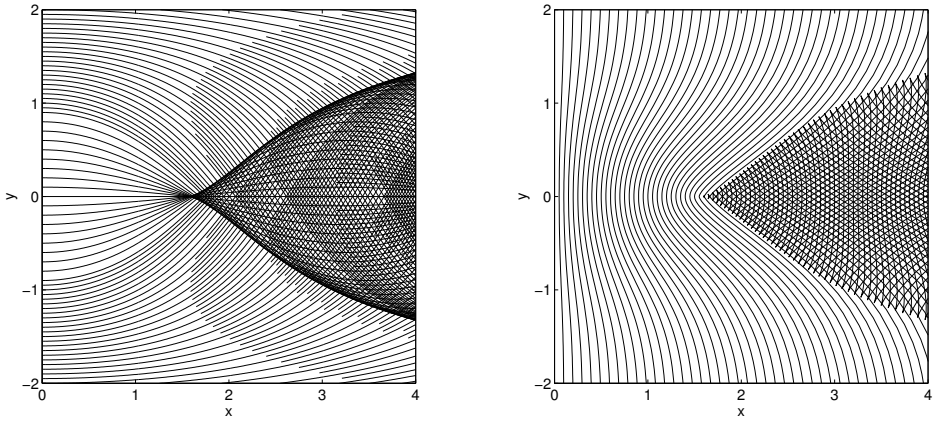
In this section, we perform two numerical tests and employ the wavefront method described in Section 3 to compute the high frequency wave fields. In both tests, we consider a rectangular computational domain $D = [0, 4] \times [-2, 2]$. The wave field is generated by a plane wave which propagates into the domain from the left boundary on the y-axis and orthogonal to the boundary. Consequently, we obtain the initial conditions for the ray tracing equations (10) as $\boldsymbol{x}_0(s) = (0, s)$ and $\theta_0(s) = 0$. The plane wave is refracted as it propagates through the domain with a variable speed of propagation and form different types of caustics, including cusps and folds. In all computations, we employ the fourth-order Runge-Kutta method for solving the ODE systems (10) and (20). We use a fixed initial spacing of Gaussian beams $h = 0.005$, which gives $N_{GB} = 1 + 4/h = 801$ beams. Unless stated otherwise, we use a fixed set of initial parameters $(Q_0, P_0) = (1, i)$.

4.1. Numerical test 1

We consider the following speed of propagation

$$c(x, y) = \frac{1}{1 + e^{-y^2}}, \qquad (x, y) \in D.$$

The plane wave propagating from the left boundary is refracted inside the domain, and a cusp caustic and two fold caustics are formed. Figure 4 shows the central rays of the Gaussian beams and the corresponding wavefronts.



(a) Central rays (b) Wavefronts

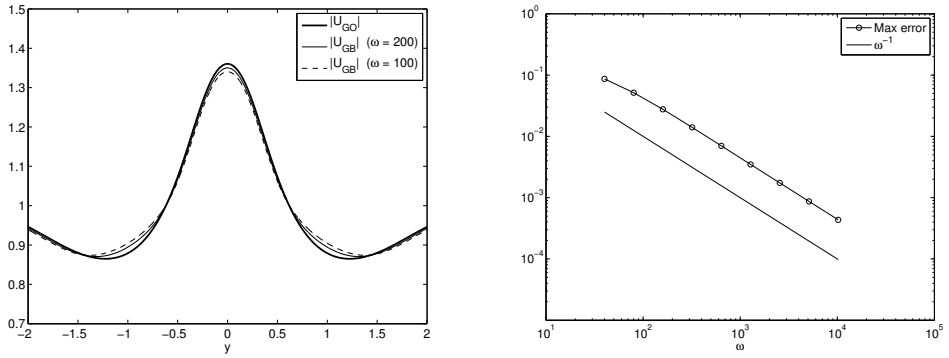
Figure 4: Test 1. Central rays and wavefronts generated by a plane wave propagating into the domain from the left boundary. The wave field is refracted inside the domain and forms a cusp and two fold caustics.

The modulus of the total wave field computed by the wavefront method along the line $x = 1$ (before caustics) is shown in Figure 5(a) for two different frequencies $\omega = 100, 200$. We use the standard GO solution, which has an error of order $\mathcal{O}(\omega^{-1})$, as a reference solution. As can be seen, the solution obtained by the Gaussian beam method converges to the solution obtained by geometrical optics as the frequency increases. Figure 5(b) shows the maximum pointwise difference between the Gaussian beam solution and the geometrical optics solution. The difference is proportional to ω^{-1} and agrees with the convergence rate obtained in [17].

As mentioned in Section 3, in order to find the solution along the line $x = 1$, we can interpolate the solution along the wavefronts down on the line $x = 1$. However, here, we employ a paraxial formulation of the governing ODEs (10) and (20) in which the variable x is a time-like independent variable. In this setting, by choosing $x = 1$ as the final value of the independent variable, we obtain the solution along the line $x = 1$ without requiring interpolation.

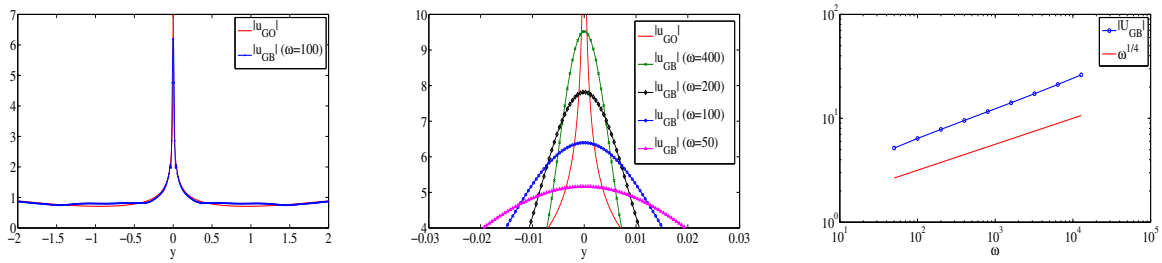
Figure 6(a) shows the total wave field along the line $x = 1.572$ where a cusp caustic is formed at $y = 0$. A zoomed view at the caustic is shown in Figure 6(b). Unlike the amplitude of the geometrical optics solution, which is unbounded at the caustic, the amplitude of the Gaussian beam solution is bounded and increases as the frequency increases. The rate of increase is shown in Figure 6(c) and agrees with the Maslov theory, which predicts $|u| = \mathcal{O}(\omega^{1/4})$ at a cusp caustic. See e.g [43].

Figure 7 shows the total wave field along the line $x = 2.5$ (after the cusp where two folds are formed). Note that in between the fold caustics, there are multiple arrival times, and the amplitude of the wave field is very oscillatory. A zoomed view around the fold caustic is shown in Figure 7(b). While the amplitude of the geometrical optics solution is unbounded



(a) (b)

Figure 5: Test 1. (a) Magnitude of the solution obtained by the wavefront Gaussian beam method with different frequencies and by geometrical optics at $x = 1$. (b) The logarithmic scale of the maximum pointwise difference between the Gaussian beam solutions and the geometrical optics solution. The difference is of order $\mathcal{O}(\omega^{-1})$.

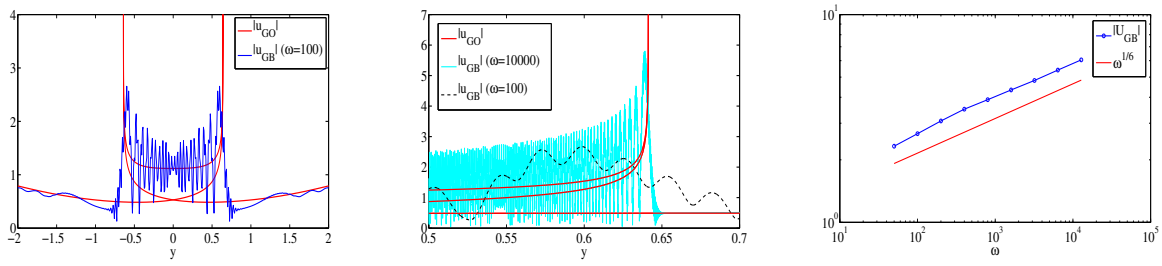


(a) (b) (c)

Figure 6: Test 1. (a) Absolute value of the wave field $|u|$, computed by GO and GB methods, along the line $x = 1.572$. A cusp caustic is formed at $(x, y) = (1.572, 0)$. (b) A zoomed view of the solution magnitude close to the cusp caustic. While the amplitude of the GO solution is unbounded at the caustic, the GB solutions are bounded and increase as the frequency increases. (c) Rate of increase agrees with Maslov theory.

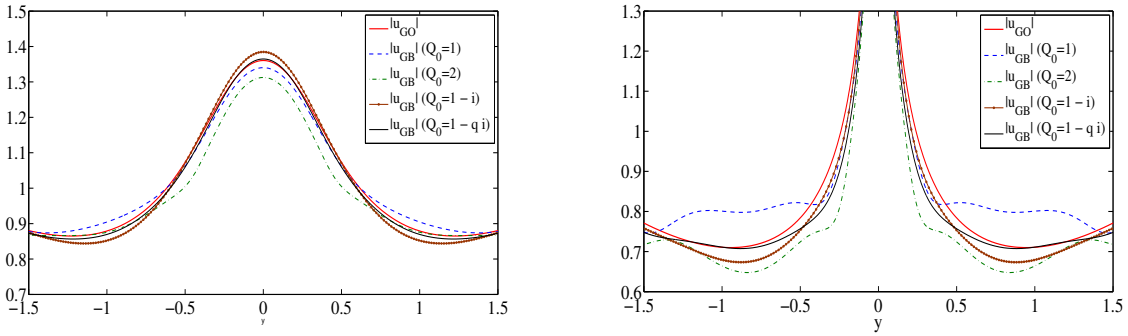
at the caustic, the amplitude of the Gaussian beam solution is bounded and increases as the frequency increases. The rate of increase is shown in Figure 7(c) and agrees with the Maslov theory, which predicts $|u| = \mathcal{O}(\omega^{1/6})$ at a fold caustic. See e.g [43].

Optimization. As mentioned in Section 2.5, a main advantage of the proposed wavefront algorithm is that by using the canonical solutions (Q_I, P_I) and (Q_{II}, P_{II}), we can compute the solution at different points of the domain with different initial data (Q_0, P_0) at no extra cost. It provides a simple and fast mechanism for optimizing the solution. In order to verify this, we compute and plot the magnitude of the solution along two different lines $x = 1$ and $x = 1.572$ in the following way, see Figure 8. First, we use different fixed values $Q_0 = 1, 2, 1-i$ with $P_0 = i$ for the initial data and calculate the corresponding solutions. Next, we use the approach discussed in Section 2.5 and obtain optimal complex-valued initial data, which generate beams with small widths along the lines $x = 1$ and $x = 1.572$. In particular, we



(a) (b) (c)

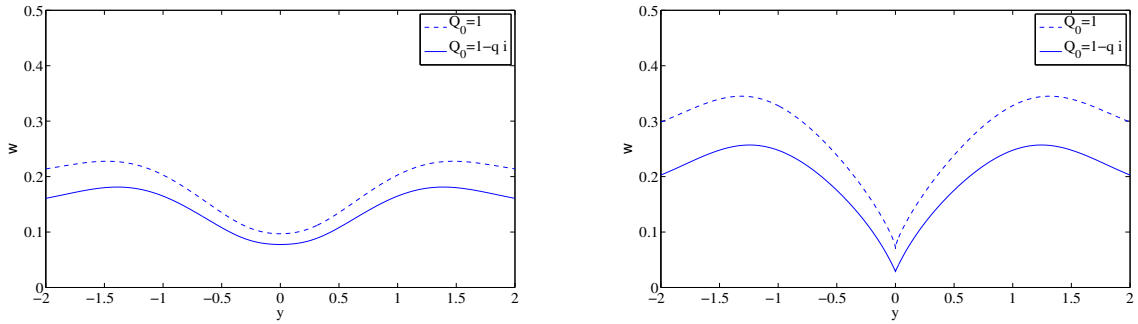
Figure 7: Test 1. (a) Absolute value of the wave field $|u|$, computed by GO and GB methods, along the line $x = 2.5$. Two fold caustics are formed at $y = \pm 0.641$ along this line. (b) A zoomed view of the solution magnitude close to the fold caustic at $(x, y) = (2.5, 0.641)$. While the amplitude of the GO solution is unbounded at the caustic, the GB solutions are bounded and increase as the frequency increases. (c) Rate of increase agrees with Maslov theory.



(a) $x = 1$, $q = 0.78\, Q_{II}/Q_I$ (b) $x = 1.572$, $q = 0.6\, Q_{II}/Q_I$

Figure 8: Test 1. The magnitude of the Gaussian beam solution with different initial data along the lines $x = 1$ (left) and $x = 1.572$ (right). We select the frequency $\omega = 100$.

select $Q_0 = 1 - q\, i$ and $P_0 = i$, where $q = 0.78\, Q_{II}/Q_I$ along $x = 1$ and $q = 0.6\, Q_{II}/Q_I$ along $x = 1.572$. We note that Q_0 is no longer fixed and varies for different points along the lines $x = 1$ and $x = 1.572$, because the canonical solutions vary. The optimally chosen complex-valued parameters generate narrower beams compared to those generated by the fixed parameters. Figure 9 shows, for instance, the half-width of the beams for the fixed parameter $Q_0 = 1$ and the optimal parameter $Q_0 = 1 - q\, i$. As shown in Figure 8, by choosing different initial data at different points, it is possible to improve the solution. For instance, the difference compared to geometrical optics, measured in relative L^2-norm at $x = 1.572$ and along $y \in [-1.4, -0.6]$ is reduced from 10% when $Q_0 = 1$ to less than 1% by choosing the optimal initial parameters.



(a) $x = 1$, $q = 0.78\, Q_{II}/Q_I$ (b) $x = 1.572$, $q = 0.6\, Q_{II}/Q_I$

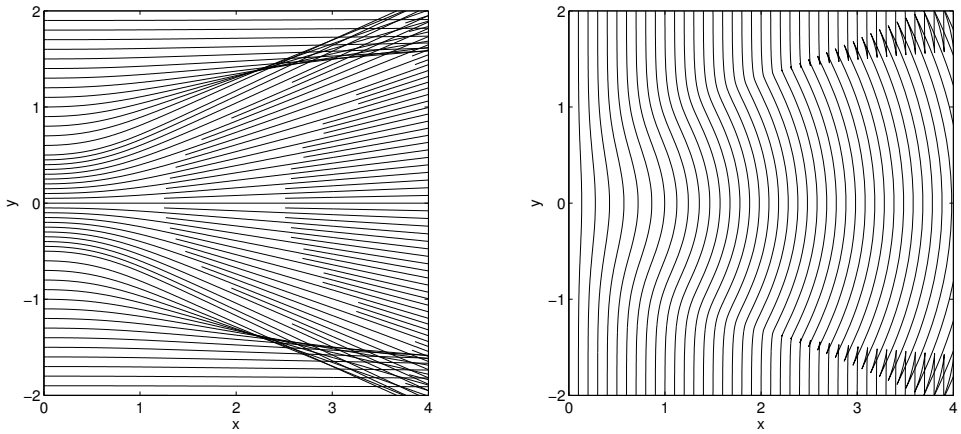
Figure 9: Test 1. The half-width of the Gaussian beams solution with different initial data along the lines $x = 1$ and $x = 1.572$. We select the frequency $\omega = 100$.

4.2. Numerical test 2

As a second test, we choose the following speed of propagation

$$c(x, y) = 1 + 0.5\, e^{-2\,((x-0.5)^2+y^2)}, \qquad (x, y) \in D.$$

In this case, the plane wave forms two cusp caustics followed by four fold caustics. Figure 10 shows the central rays of the Gaussian beams and the corresponding wavefronts.

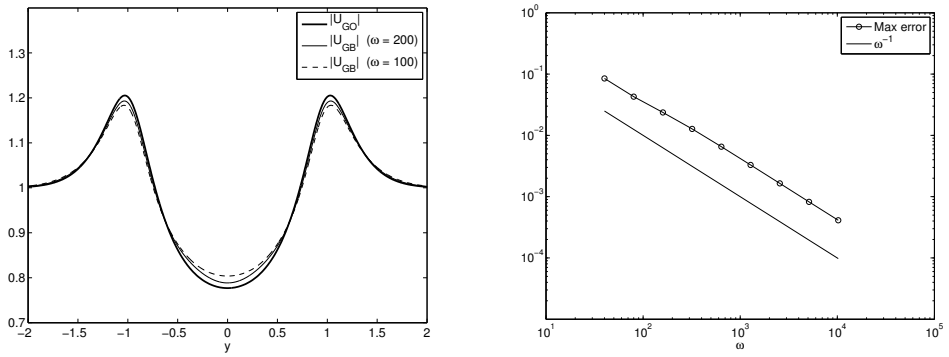


(a) Central rays (b) Wavefronts

Figure 10: Test 2. Central rays and wavefronts generated by a plane wave propagating into the domain from the left boundary. The wave field is refracted inside the domain and forms two cusp and four fold caustics.

The total wave field for different frequencies along the line $x = 1$ and the maximum pointwise difference between the Gaussian beam solution and the geometrical optics solution are shown in Figure 11. The difference is proportional to ω^{-1}, as expected.

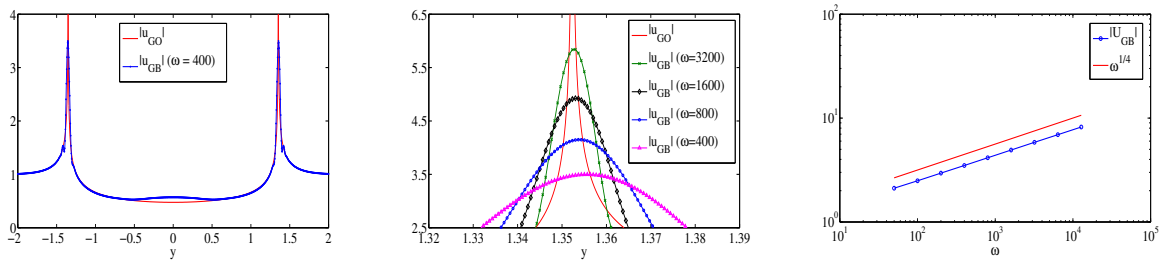
Figure 12(a) shows the total wave field along the line $x = 2.125$, where two cusp caustics are formed at $y = \pm 1.352$. A zoomed view close to the cusp caustic at $(x, y) = (2.125, 1.352)$



(a) (b)

Figure 11: Test 2. (a) Magnitude of the solution obtained by the wavefront Gaussian beam method with different frequencies and by geometrical optics at $x = 1$. (b) The logarithmic scale of the maximum pointwise difference between the Gaussian beam solutions and the geometrical optics solution. The difference is of order $\mathcal{O}(\omega^{-1})$.

is shown in Figure 12(b), and the rate of increase of the Gaussian beam solutions as the frequency increases is shown in Figure 12(c). As it can be seen, we observe that $|u| = \mathcal{O}(\omega^{1/4})$, which is in agreement with the Maslov theory.

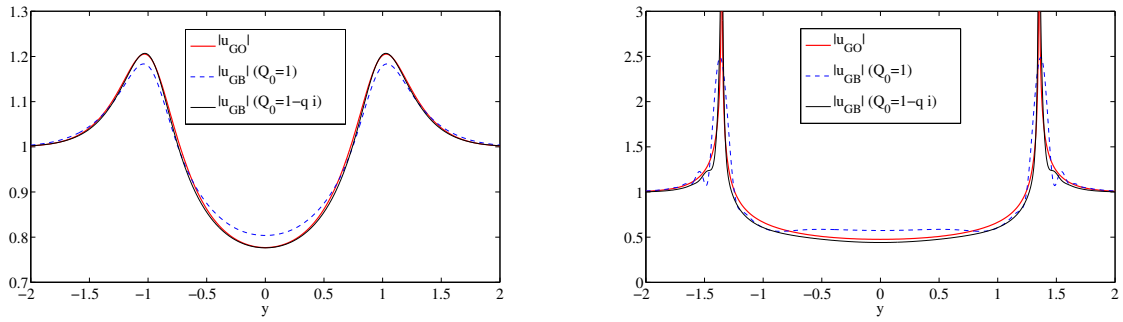


(a) (b) (c)

Figure 12: Test 2. (a) Absolute value of the wave field $|u|$ along the line $x = 2.125$. Two cusp caustics are formed at $y = \pm 1.352$. (b) A zoomed view of the solution magnitude close to the cusp caustic at $(x, y) = (2.125, 1.352)$. While the amplitude of the GO solution is unbounded at the caustic, the GB solutions are bounded and increase as the frequency increases. (b) Rate of increase agrees with Maslov theory.

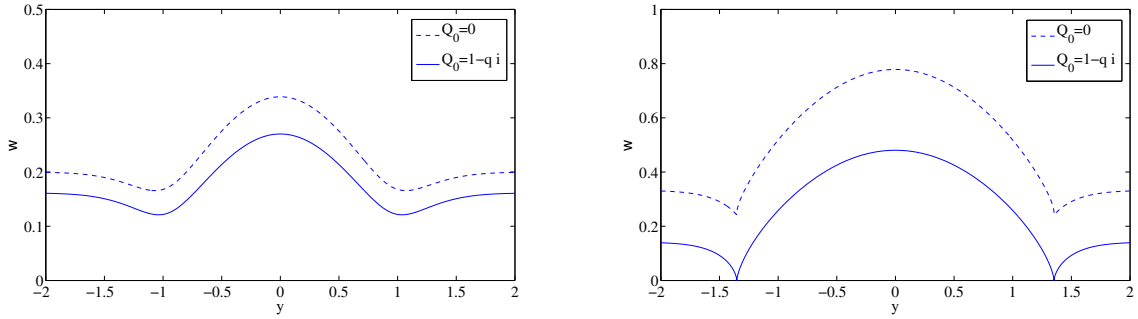
Optimization. Similarly to numerical test 1, we use complex-valued initial data, which generate beams with small widths along the lines $x = 1$ and $x = 2.125$. We compute and plot the magnitude of the solution along these two lines in the following way, see Figure 13. First, we use the fixed values $Q_0 = 1, P_0 = i$ for the initial data and calculate the solution. Next, we use the approach, as discussed in Section 2.5, and obtain complex-valued initial data $Q_0 = 1 - q\,i$ and $P_0 = i$, where $q = 0.45\,Q_{II}/Q_I$ along $x = 1$ and $q = Q_{II}/Q_I$ along $x = 2.125$. These complex-valued parameters generate narrower beams compared to those generated by the fixed parameter $Q_0 = 1$, see Figure 14. As shown in Figure 13, by choosing different initial data at different points, it is possible to improve the solution. For instance, the difference compared to geometrical optics, measured in relative L^2-norm at $x = 2.125$

and along $y \in [-1, 1]$ is reduced from 14% to 7% by choosing the optimal initial parameters.



(a) $x = 1$, $q = 0.45\, Q_{II}/Q_I$ (b) $x = 2.125$, $q = Q_{II}/Q_I$

Figure 13: Test 2. The magnitude of the Gaussian beam solution with different initial data along the lines $x = 1$ (left) and $x = 2.125$ (right). We select the frequency $\omega = 100$.



(a) $x = 1$, $q = 0.45\, Q_{II}/Q_I$ (b) $x = 2.125$, $q = Q_{II}/Q_I$

Figure 14: Test 2. The half-width of the Gaussian beams solution with different initial data along the lines $x = 1$ and $x = 2.125$. We select the frequency $\omega = 100$.

5. Conclusion

We have proposed a novel wavefront-based Gaussian beam method for computing high frequency wave propagation problems. The method tracks a front of two canonical beams with two particular initial values for width and curvature. Using the two sets of canonical solutions along the front, we can efficiently recreate Gaussian beams with any arbitrary initial data on the initial front. This provides a simple mechanism to include a variety of optimization processes, including error minimization or beam width minimization, for a posteriori selection of optimal beams' initial parameters, which give more accurate solutions. The performed numerical examples illustrate the accuracy, efficiency, and the flexibility of the method in selecting initial parameters.

Since the geometrical optics solution can be recovered by the first set of canonical solutions, it is possible to design a hybrid method which switches between the geometrical

optics and Gaussian beam solutions smoothly. Such a hybrid method would substantially increase the efficiency of the Gaussian beam wavefront method, because away from caustics, the computation of Gaussian beams, which requires a post-processing step, is not needed. The construction of a hybrid algorithm based on an optimal selection of initial parameters is the subject of our current work and will be presented elsewhere.

References

[1] B. Engquist, O. Runborg, Computational high frequency wave propagation, Acta Numerica 12 (2003) 181–266.

[2] O. Runborg, Mathematical models and numerical methods for high frequency waves, Commun. Comput. Phys. 2 (2007) 827–880.

[3] M. M. Popov., A new method of computation of wave fields using Gaussian beams, Wave Motion 4 (1982) 85–97.

[4] V. M. Babič, T. F. Pankratova, On discontinuities of Green's function of the wave equation with variable coefficient, Problems of Mathematical Physics Leningrad University Press 6 (1973) 9–27, (in Russian).

[5] A. P. Katchalov, M. M. Popov, Application of the method of summation of Gaussian beams for calculation of high-frequency wave fields, Sov. Phys. Dokl. 26 (1981) 604–606.

[6] V. C. Červený, M. M. Popov, I. Pšenčík, Computation of wave fields in inhomogeneous media — Gaussian beam approach, Geophys. J. R. Astr. Soc. 70 (1982) 109–128.

[7] L. Klimeš, Expansion of a high-frequency time-harmonic wavefield given on an initial surface into Gaussian beams, Geophys. J. R. astr. Soc. 79 (1984) 105–118.

[8] E. J. Heller, Frozen Gaussians: a very simple semiclassical approximation, J. Chem. Phys. 76 (1981) 2923–2931.

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

[10] N. R. Hill, Gaussian beam migration, Geophysics 55 (1990) 1416–1428.

[11] N. R. Hill, Prestack Gaussian beam depth migration, Geophysics 66 (2001) 1240–1250.

[12] J. Ralston, Gaussian beams and the propagation of singularities, in: W. Littman (Ed.), Studies in partial differential equations, volume 23 of MAA Stud. Math., Vol. 23, Math. Assoc. America, Washington, DC, 1982, pp. 206–248.

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

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

[15] J. Lu, X. Yang, Convergence of frozen gaussian approximation for high frequency wave propagation, Comm. Pure Appl. Math. 65 (2012) 759–789.

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

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

[18] H. Liu, O. Runborg, J. Ralston, N. Tanushev, Gaussian beam methods for the Helmholtz equation, SIAM J. Appl. Math. 74 (3) (2014) 771–793.

[19] N. M. Tanushev, J. Qian, J. V. Ralston, Mountain waves and Gaussian beams, Multiscale Model. Simul. 6 (2007) 688–709.

[20] E. Faou, C. Lubich, A poisson integrator for gaussian wavepacket dynamics, Computing and Visualization in Science 9 (2006) 45–55.

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

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

[23] S. Leung, J. Qian, R. Burridge, Eulerian Gaussian beams for high frequency wave propagation, Geophysics 72 (2007) SM61–SM76.

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

[25] S. Jin, H. Wu, X. Yang, Z. Y. Huang, Bloch decomposition-based Gaussian beam method for the Schrödinger equation with periodic potentials, J. Comput. Phys. 229 (2010) 4869–4883.

[26] S. Leung, J. Qian, Eulerian Gaussian beams for Schrödinger equations in the semi-classical regime, J. Comput. Phys. 228 (2009) 2951–2977.

[27] S. Jin, P. Markowich, C. Sparber, Mathematical and computational models for semi-classical Schrödinger equations, Acta Numerica 21 (2012) 1–89.

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

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

[30] V. Vinje, E. Iversen, H. Gjoystdal, Traveltime and amplitude estimation using wavefront construction, In Eur. Ass. Expl. Geophys. (1992) 504–505.

[31] V. Vinje, E. Iversen, H. Gjoystdal, Traveltime and amplitude estimation using wavefront construction, Geophysics 58 (1993) 1157–1166.

[32] J. Qian, L. Ying, Fast Gaussian wavepacket transforms and Gaussian beams for the Schrödinger equation, J. Comput. Phys. 229 (2010) 7848–7873.

[33] B. S. White, A. Norris, A. Bayliss, R. Burridge, Some remarks on the Gaussian beam summation method, Geophys. J. R. astr. Soc. 89 (1987) 579–636.

[34] Y. Sun, Computaton of 2d multiple arrival traveltime fields by an interpolative shooting method, Soc. Expl. Geophys., Extended Abstracts (1992) 1320–1323.

[35] S. Leung, H. Zhao, Gaussian beam summation for diffraction in inhomogeneous media based on the grid based particle method, Commun. Comput. Phys. 8 (2010) 758–796.

[36] O. Runborg, Fast interface tracking via a multiresolution representation of curves and surfaces, Commun. Math. Sci. 7 (2009) 365–398.

[37] O. Runborg, Analysis of high order fast interface tracking methods, Numer. Math. 128 (2014) 339–375.

[38] M. Motamed, O. Runborg, Asymptotic approximations of high frequency wave propagation problems, in: B. Engquist, A. Fokas, E. Hairer, A. Iserles (Eds.), Highly Oscillatory Problems, volume 366 of London Mathematical Society Lecture Note Serie, Cambridge University Press, 2009, pp. 72–97.

[39] D. Ludwig, Uniform asymptotic expansions at a caustic, Comm. Pure Appl. Math. 19 (1966) 215–250.

[40] A. P. Katchalov, M. M. Popov, Application of the Gaussian beam method to elasticity theory, Geophys. J. R. Astr. Soc. 81 (1985) 205–214.

[41] L. Klimeš, Optimization of the shape of Gaussian beams of a fixed length, Stud. Geophys. Geod. 33 (1989) 146–163.

[42] G. Lambaré, P. S. Lucio, A. Hanyga, Two-dimensional multivalued traveltime and amplitude maps by uniform sampling of ray field, Geophys. J. Int. 125 (1996) 584–598.

[43] M. E. Taylor, Partial Differential Equations I: Basic Theory, volume 115 of Applied Mathematical Sciences, Springer, New York, 2010.