Numerical microlocal analysis of harmonic wavefields

Jean-David Benamou Francis Collino Olof Runborg

March 15, 2004

Abstract

We present and test a numerical method which, given an analytical or numerical solution of the Helmholtz equation in a neighborhood of a fixed observation point and assuming that the geometrical optics approximation is relevant, determines at this point the number crossing rays and computes their directions and associated complex amplitudes.
Keywords : geometrical optics, microlocal analysis, Helmholtz equation

1 Introduction

We start with the derivation of the geometrical optics (GO) model (see [7] for more details including the geometrical theory of diffraction (GTD)). Let $u_k(x)$ be the solution of the Helmholtz equation

$$\Delta u_k + k^2\eta^2 u_k = 0, \tag{1}$$

supplemented by suitable boundary and radiation conditions. The coefficient $\eta = \eta(x)$ is the index of refraction and $k = 2\pi/\lambda_0$ is the wavenumber where λ_0 is the reference wavelength. In its simplest form, geometrical optics relies on the assumption that the complex valued solution $u_k(x)$ can be approximated, asymptotically in k, by the "ansatz"

$$u_k(x) \simeq A(x)e^{ik\phi(x)}, \tag{2}$$

where the amplitude $A(x)$ and the phase $\phi(x)$ are frequency independent real valued smooth solutions of the Eikonal/transport GO system of equations,

$$|\nabla\phi| = \eta, \qquad 2\nabla\phi\cdot\nabla A + A\Delta\phi = 0. \tag{3}$$

We will always assume that the amplitude $A(x)$ is positive. This analysis relies on the following intuitive idea : when the wavelength is much smaller than the scale of variations of the index of refraction $\eta(x)$ characterizing the medium, the solution locally behaves as an elementary plane wave. Indeed, as A and ϕ are assumed to be smooth, a first order approximation around x_0 gives the local plane wave approximation

$$u_k(x) \simeq B(x_0) e^{ik(x-x_0)\cdot\nabla\phi(x_0)} \tag{4}$$

where we denote $B(x_0) = A(x_0)e^{ik\phi(x_0)}$ the "complex amplitude".

Classically, system (3) is solved by the method of characteristics, called "rays" in this context; the rays are the integral curves of the vector field $\nabla\phi$ and thus follow the "local" plane wave directions $\nabla\phi(x_0)$. Note that in the GTD the A coefficient may depend on k (we treat such a case in Section 3.5.3).

The ray field is of course computed independently of $\nabla\phi$ and general solutions may exhibit an arbitrary number of crossing rays (due to reflection, diffraction, fold or collapse phenomena). The ansatz (2) is then not relevant and more sophisticated mathematics are needed (see [14] for instance). Away from caustics and focus points the situation simplifies, and the relevant asymptotic theory consists in locally approximating the solution as a superposition of a finite number N of elementary ansatz of type (2)

$$u_k(x) \simeq \sum_{n=1}^{N} A_n(x) e^{ik\phi_n(x)}. \tag{5}$$

The number of elementary contributions and their coefficients ϕ_n and A_n depend on the number of rays crossing at x and their associated phases and amplitudes. See [3, 16] for a short presentation of geometrical optics and a discussion about the associated concept of multi-valued solutions to (3).

In this paper we will use the geometrical optics approximation (5) in its plane wave approximate form (4)

$$u_k(x) \simeq \sum_{n=1}^{N} B_n(x_0) e^{ik(x-x_0)\cdot\nabla\phi_n(x_0)}, \tag{6}$$

where $B_n(x_0) = A_n(x)e^{ik\phi_n(x_0)}$, and consider the following inverse problem :

> Given an analytical or numerical solution $u_k(x)$ in a neighborhood of a fixed observation point x_0, determine the number of rays N crossing at x_0 and compute the GO quantities $(B_n, \nabla\phi_n)$ for $n = 1, \dots, N$.

As an illustration, the left part of Figure 1 shows the modulus of a Helmholtz solution made up of a sum of two circular waves in homogeneous space. The right part of the figure shows the associated ray directions $\nabla\phi_n$ at a finite number of observation points forming a grid : each point is passed by two rays issued from the sources.

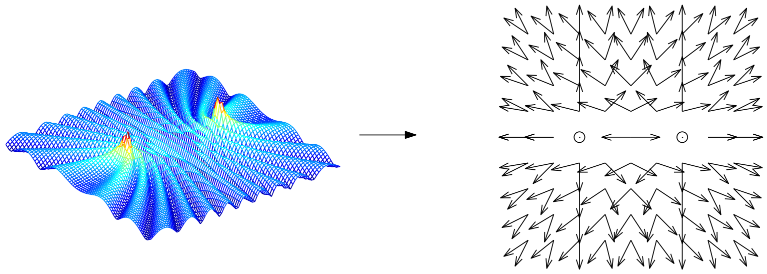


Figure 1: Translating a Helmholtz solution to rays.

The general motivation for this problem is to reduce the complexity, and cost, of many numerical problems by being able to move from a detailed description to a description only involving the smooth functions $A_n(x)$ and $\phi_n(x)$, which can be represented by a fixed number of unknowns, independent of the frequency, and which can be processed via the "coarse scale" geometrical optics equations (3). The applications we have in mind include:

- Hybrid solvers in which the Helmholtz equation is solved in some (complicated) parts of the domain and geometrical optics are used in the other parts to reduce computational costs. This technique is often used in computational electromagnetics (CEM), see e.g. [8, 34, 29, 21, 2, 20] amongst others. The same idea applies to domain decomposition methods where there is a coupling between domains in which the Helmholtz equation is solved and domains where ray tracing is used. Another potential application would be to use this method to construct absorbing boundary conditions for the Helmholtz equation.

- Different methods dealing with the construction of Galerkin methods for wave type equations based on "GO" basis functions [18, 10, 1, 13]. The possibility to analyze a local high frequency field could be helpful either for error estimations or even the construction of ad-hoc preconditioners.

- The method can also provide useful information on the relevance of the GO approximation of the solution. As the GO solution is frequency independent (except for diffraction phenomena), the quality of the GO output of our algorithm increases with the frequency and therefore it can detect the frequency threshold where one can use the GO quantities to extrapolate the Helmholtz solution further, c.f. comment in Section 3.4.

- As a restriction/compression operator in coarse timestepper based or heterogeneous multiscale methods for wave problems [15, 22]. In these methods one needs to explicitly convert local detailed descriptions into a global coarse description. The opposite conversion, lifting/reconstruction, can be trivially performed by just evaluating the expression in (6).

- When considering electromagnetic fields generated by one or more antenna, this problem is called DOA (Direction Of Arrival) estimation in the signal processing community. It is either solved by parametric fitting assuming some knowledge of the impinging signal (MUSIC algorithm [31]) or by exploiting the phase shift in the propagation model in the case of a linear array of antennas (ESPRIT algorithm [30]).

One possible solution to our inverse problem when the observation point is far from the heterogeneous zone is to use a far-field approximation. Let Γ be a closed curve including either a scatterer or some heterogeneities. When Γ is small enough compared to the wavelength, the interior domain can be interpreted as one single point source. GO rays can be considered as flowing isotropically from the scattering zone and the far field prescribes the amplitude as a function of the directions. This approach is only accurate quite far from Γ though. The distance D and the scatter size d must satisfy $kd^2 \ll D$ and $d \ll D$, and it cannot capture crossing rays in the vicinity of Γ.

Another natural idea is to analyze the restriction of u_k on a surface or an interface, in terms of some functions that have a simple ray interpretation. A Fourier transform for instance will directly provide a plane wave analysis (possibly after pre-processing). In GO, a plane wave is just a family of parallel rays that sweep the domain in a prescribed direction. A more sophisticated version of this approach uses a decomposition in Gaussian beams [27] [28].

One could also consider this problem using Huygen's principle, which is perhaps the most common approach in CEM. It can be understood using the integral form of the solution,

$$u_k(x) = \int_\Gamma \Phi_k(x,y)\frac{\partial u_k}{\partial \nu}(y) - \frac{\partial \Phi_k(x,y)}{\partial \nu(y)} u_k(y)\, ds(y), \tag{7}$$

outside any closed curve Γ. Typically, the curve is also in this case the boundary of a scatterer or a penetrable heterogeneous local zone. The function $\Phi_k(x,y)$ is the free space Green function which can be interpreted as a circular (spherical in 3-D) wave emanating from the point y,

$$\Phi_k(x,y) \simeq \frac{e^{ik|x-y|}}{|x-y|^{\frac{d-1}{2}}}, \qquad |x-y| \gg 1,$$

in dimension d. The normal derivative of $\Phi_k(x,y)$ behaves similarly. Upon approximating the integral (7) with a sum, we obtain a solution of the form (5). We can think of this as splitting the curve Γ into small pieces, each considered as a secondary source from which rays are propagated in all directions.

One must realize though that these methods do not directly give the high frequency approximation of the solution in the sense of (5) but rather the high frequency approximation of an "interpretation" of the solution in terms of secondary sources. They rely on cancellation effects between nearby sources and therefore a significant number of rays must be used to get an accurate approximation, the number increasing with the frequency. For example, a plane wave could be approximated by a sum of circular waves, but for fixed accuracy the number N of such waves would have to be of the order k. In contrast, for the high frequency approximation a plane wave would have $N = 1$ independent of k. Note also that the integral form (7) and the far field approximation are only valid where the medium is homogeneous.

In this paper, we propose an algorithm that really achieves the above local inverse problem and the cost of the algorithm is frequency independent. By "local" we mean that we only need the Helmholtz solution in a neighborhood of the "observation" point for a (sufficiently large) fixed k to obtain the "exact" GO asymptotic interpretation of the solution (i.e. the GO solution as if it had been globally computed using an asymptotic GO model). Our method works both for homogeneous and inhomogeneous problems. It essentially relies on the study of the restriction of the solution to a small circle around each point we want to analyze. More precisely, if α is some given positive number, we construct the function

$$U_\alpha : \hat{s} \quad \longrightarrow \quad u_k\left(x_0 + \frac{\alpha}{k\eta(x_0)}\hat{s}\right),$$

where $\hat{s}$ runs over the circle (2-D case) or the sphere (3-D case) of radius one. In other words, we will analyze the wave at some points on a circle (or a sphere) centered at x_0 whose radius is $\alpha/2\pi$ times the wavelength measured at x_0. Analyzing locally a function along some given directions is known as a microlocal analysis (in a broader sense however than what is called precisely microlocal analysis in the mathematical analysis field). Our goal therefore is to synthesize all significant microlocal directions, whence the title of this paper. From the knowledge of the function U_α, we will show that it is possible to recover numerically some information about the number of rays crossing at x_0 as well as their complex amplitudes and directions.

Outline

After stating the "numerical microlocal" formulation of the problem and the filtering procedure based on the Jacobi–Anger formula in Section 2, we detail the 2-D algorithm and present numerical results in Section 3. Section 4 presents the 3-D extension.

2 Numerical microlocal analysis

2.1 General setting

In this section, we assume that the values of the solution around the point at which we look at are directly available, either in analytic form or after some interpolation procedure if the solution has been computed using, for instance, a finite element or finite difference method. A separate analysis based on Herglotz waves is being investigated when the solution comes from an integral equation.

Let x_0 be some point in the space and $u_k(x)$ a solution to the Helmholtz equation with wave number k in the neighborhood of x_0. We assume that there exists some integer N and some phase functions and amplitudes, $\phi_n(x)$ and $A_n(x)$, $n = 1, \dots, N$, such that

$$u_k(x) \simeq \sum_{n=1}^{N} A_n(x) e^{ik\phi_n(x)}, \tag{8}$$

when $|x - x_0|$ is small. The phase functions are assumed to satisfy the Eikonal equation

$$|\nabla \phi_n(x)| = \eta(x),$$

in the domain where this n-th branch of the phase family exists and contributes to the global GO solution.

Henceforth, we will denote by $\hat{d}_n(x)$ the direction of propagation of the rays

$$\nabla \phi_n(x) = \eta(x)\hat{d}_n(x).$$

We consider the Helmholtz solution for wave number k on a circle or a sphere of radius $\alpha/k\eta(x_0)$ around the point x_0. Thus, the parameter α is the radius scaled by the wavelength divided by 2π. We define

$$U_\alpha(\hat{s}) := u_k\left(x_0 + \frac{\alpha}{k\eta(x_0)}\hat{s}\right). \tag{9}$$

This is a function whose argument varies on the unit circle or the unit sphere. The function U_α also depends on k and on x_0 but since these parameters will be kept fixed in the remaining discussion, we do not make the dependence explicit in the notation to enhance readability. Using the Taylor expansions,

$$\begin{aligned}\phi_n(x) &= \phi_n(x_0) + \nabla\phi_n(x_0)\cdot(x-x_0) + \dots \\ &= \phi_n(x_0) + \eta(x_0)\hat{d}_n(x_0)\cdot(x-x_0) + \dots \\ A_n(x) &= A_n(x_0) + \dots,\end{aligned} \tag{10}$$

we see that by (8), we have

$$U_\alpha(\hat{s}) \simeq \sum_{n=1}^{N} A_n(x_0)\, e^{ik\phi_n(x_0)}\, e^{i\alpha\,\hat{s}\cdot\hat{d}_n(x_0)}, \tag{11}$$

which will be processed under the following form:

$$U_\alpha^{\mathrm{ray}}(\hat{s}) := \sum_{n=1}^{N} B_n(x_0)\, e^{i\alpha\,\hat{s}\cdot\hat{d}_n(x_0)}, \tag{12}$$

where $B_n(x_0) = A_n(x_0)\, e^{ik\phi_n(x_0)}$ is the "complex amplitude" at x_0. Note that for large k, the function $B_n(x_0)$ is a highly oscillating function of x_0 and it may be difficult to recover $\phi_n(x_0)$ from a numerical approximation. Its modulus, on the other hand, is smooth and equal to $A_n(x_0)$.

We now investigate the recovery of the ray directions from the knowledge of the function $U_\alpha(\hat{s})$. The idea is to remove the effect of the exponential factor by filtering and to calculate a function $\beta_\alpha(\hat{s})$ that has distinct peaks in the ray directions.

In the following sections we will drop the explicit dependence on x_0 also in the notation for $\hat{d}_n$ and B_n.

2.2 2-D case

We introduce the angle notation : $\theta_n = \theta(\hat{d}_n)$ and $\theta(\hat{s})$ such that

$$\hat{s} = (\cos\theta(\hat{s}), \sin\theta(\hat{s})), \quad \hat{d}_n = (\cos\theta_n, \sin\theta_n).$$

Our basic tool is the 2-D Jacobi–Anger expansion (see [12] page 66) :

$$e^{i\alpha\, \hat{s}\cdot\hat{d}_n} = e^{i\alpha\cos(\theta_n - \theta(\hat{s}))} = \sum_{\ell=-\infty}^{\infty} i^\ell J_\ell(\alpha) e^{-i\ell(\theta_n - \theta(\hat{s}))}, \tag{13}$$

where $J_\ell(\alpha)$ is the Bessel function of order ℓ. Inserting the Jacobi–Anger expansion into (12), we get an expression for $U_\alpha^{\mathrm{ray}}(\hat{s})$,

$$U_\alpha^{\mathrm{ray}}(\hat{s}) \simeq \sum_{\ell=-\infty}^{\infty} i^\ell J_\ell(\alpha) \left(\sum_{n=1}^{N} B_n e^{i\ell(\theta(\hat{s}) - \theta_n)} \right). \tag{14}$$

We recall that α is the radius parameter scaled by the wavelength divided by 2π.

We need to truncate the series not only to keep the computational cost down, but also to avoid numerical instabilities in the dividing operator, (17) below, when the terms in the series become small. In fact, asymptotic analysis of the Bessel function, [12], reveals that when α is held fixed, $J_\ell(\alpha)$ goes to zero more than exponentially fast with $|\ell|$ for large enough $|\ell|$. Indeed we have from [12]

$$\begin{array}{l} J_\ell(\alpha) = \frac{\alpha^\ell}{2^\ell \ell!}(1 + O(\frac{1}{\ell})), \quad \ell > 0, \\ J_{-\ell}(\alpha) = (-1)^\ell J_\ell(\alpha). \end{array} \tag{15}$$

See Figure 2 for an illustration.

When α is not too large, say $\alpha < 100$, the series can be truncated at a threshold $|\ell| \leq L(\alpha)$; the following heuristic estimate of the threshold can be found in [11] :

$$L(\alpha) = \alpha + C_1 \alpha^{\frac{1}{3}} + C_2 \log(\alpha + \pi), \tag{16}$$

where $C_1 = 5$ and $C_2 = 0$ is proposed to achieve single precision accuracy (error less than 10^{-6}) and $C_1 = 10$ for double precision. This law is valid for moderate value of α. When α is larger than 30, Chew and Song, [32], give the semiempirical law $C_1 = 0$, $C_2 = 1.8 d_0^{\frac{2}{3}}$ where d_0 is the desired number of accurate digits. Other studies about the proper choice of $L(\alpha)$ can be found in [9] and [33].

We now introduce the space of Fourier coefficients

$$l^2 = \left\{ \gamma_\ell,\ -\infty < \ell < \infty,\ \sum_{\ell=-\infty}^{\infty} |\gamma_\ell|^2 < \infty \right\},$$

and let $\mathcal{F} : L^2(S^1) \to l^2$ be the Fourier transform

$$(\mathcal{F}A(\hat{s}))_\ell := \frac{1}{\sqrt{2\pi}} \int_{S^1} A(\hat{s}) e^{-i\ell\theta(\hat{s})} d\sigma(\hat{s}),$$

with inverse

$$\left(\mathcal{F}^{-1}\gamma\right)(\hat{s}) = \frac{1}{\sqrt{2\pi}} \sum_\ell \gamma_\ell e^{i\ell\theta(\hat{s})}.$$

Let $D_\alpha : l^2 \to l^2$ be the dividing operator

$$(D_\alpha \gamma)_\ell := d_\ell^\alpha\, \gamma_\ell, \qquad d_\ell^\alpha = \begin{cases} \frac{2\pi}{2L(\alpha)+1} \frac{1}{i^\ell J_\ell(\alpha)}, & |\ell| \le L(\alpha), \\ 0, & \text{otherwise}, \end{cases} \tag{17}$$

where we assume that $J_\ell(\alpha) \neq 0$ for $|\ell| \le L(\alpha)$. We then define the function

$$\beta_\alpha := \mathcal{F}^{-1} D_\alpha \mathcal{F} U_\alpha,$$

or, more explicitly,

$$\beta_\alpha(\hat{s}) = \frac{1}{\sqrt{2\pi}} \sum_{\ell=-L(\alpha)}^{L(\alpha)} d_\ell^\alpha\, (\mathcal{F}U_\alpha)_\ell\ e^{i\ell\theta(\hat{s})}. \tag{18}$$

Similarly, with $\beta_\alpha^{\text{ray}} = \mathcal{F}^{-1} D_\alpha \mathcal{F} U_\alpha^{\text{ray}}$ where U_α^{ray} is given in (12), we easily obtain

$$\beta_\alpha^{\text{ray}}(\hat{s}) = \sum_{n=1}^{N} \frac{B_n}{2L(\alpha)+1} \sum_{\ell=-L(\alpha)}^{L(\alpha)} e^{i\ell(\theta_n - \theta(\hat{s}))}$$

or,

$$\beta_\alpha^{\text{ray}}(\hat{s}) = \sum_{n=1}^{N} B_n S_\alpha(\theta(\hat{s}) - \theta_n), \qquad S_\alpha(\theta) = \frac{\sin\left([2L(\alpha)+1]\theta/2\right)}{[2L(\alpha)+1]\sin(\theta/2)}. \tag{19}$$

We close this section with a short analysis of the function $\beta_\alpha^{\rm ray}(\hat{s})$. Suppose first that only one ray exists, i.e. $N = 1$. Clearly, $S_\alpha(0) = 1$ and therefore

$$\beta_\alpha^{\rm ray}(\hat{d}_1) = B_1.$$

Moreover,

$$|S_\alpha(\theta)| \le \frac{1}{[2L(\alpha)+1]|\sin(\theta/2)|}, \qquad \theta \neq 0, \tag{20}$$

so $\beta_\alpha^{\rm ray}(\hat{s})$ goes to zero when α goes to infinity and $\hat{s} \neq \hat{d}_1$. We get a similar results when there are many rays. Let

$$\mathrm{Sep}(\theta, n) = \frac{\left|\sin\left(\frac{\theta-\theta_n}{2}\right)\right|}{|B_n|},$$

measure how well a ray in direction θ is separated from ray n, weighted by the amplitude. We then have

$$|\beta_\alpha^{\rm ray}(\hat{s}) - B_n S_\alpha(\theta(\hat{s}) - \theta_n)| \le \frac{N-1}{2L(\alpha)+1} \max_{m\neq n} \frac{1}{\mathrm{Sep}(\theta(\hat{s}), m)}, \tag{21}$$

by (19) and (20). Taking $\hat{s} = \hat{d}_n$ shows that for α large enough, we have $\beta_\alpha^{\rm ray}(\hat{d}_n) \simeq B_n$. ¿From (19) and (20) we also get that $\beta_\alpha^{\rm ray}(\hat{s}) \simeq 0$ for large α when $\hat{s} \neq \hat{d}_n$, since

$$|\beta_\alpha^{\rm ray}(\hat{s})| \le \frac{N}{2L(\alpha)+1} \max_{m} \frac{1}{\mathrm{Sep}(\theta, m)}, \qquad \hat{s} \neq \hat{d}_n,\ \forall n. \tag{22}$$

We have hence shown that for a fixed $\hat{s}$

$$\lim_{\alpha\to\infty} \beta_\alpha^{\rm ray}(\hat{s}) = \begin{cases} B_n, & \hat{s} = \hat{d}_n, \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

This analysis shows that we can expect the above filtering procedure to give as output a function $\beta_\alpha(\hat{s})$, defined on the sphere, which has sharp peaks in the directions of propagation of the rays when α is large enough.

Remark 1. *The $\beta_\alpha(\hat{s})$ function plays the same role here as the Wigner transform $W_k(x, \xi)$ ([24, 17, 4]) does in many other analyses of high frequency wave phenomena: it captures the local strength of waves propagating in different directions at an observation point $x = x_0$. When u_k is of the* *form in (5) the corresponding Wigner transform converges (weakly) with k to the Wigner measure W, given by*

$$W(x,\xi) = \sum_{n=1}^{N} A_n^2(x)\delta(\xi - \nabla\phi_n(x)), \tag{24}$$

which should be compared with (23). In fact, since $(2L(\alpha)+1)S_\alpha$ is the classical $L(\alpha)$-th Dirichlet kernel, and $L(\alpha)\to\infty$ with α, we will also have

$$(2L(\alpha)+1)\beta_\alpha^{\text{ray}}(\hat{s}) \rightharpoonup \sum_{n=1}^{N} B_n\delta(\hat{s}-\hat{d}_n),$$

as $\alpha\to\infty$. The limit of $W_k(x_0,\eta\hat{s})$, although more singular, is clearly related to the limit of $\beta_\alpha(\hat{s})$. Both give information about the local wave directions in the form of $\nabla\phi_n$ and $\hat{d}_n$. However, the limit of $\beta_\alpha(\hat{s})$ also contains informtion about the phases *at the observation point x_0 via the complex amplitudes B_n, while W only includes the amplitudes $A_n = |B_n|$. Moreover, $\beta_\alpha(\hat{s})$ is much more amenable to numerical computations: it is computed from values locally around the observation point and it converges strongly (in α) to a bounded limit by (23). The Wigner transform, on the other hand, includes a non-local Fourier transform and it converges only weakly (in k) to a measure.*

2.3 3-D case

The situation in three dimensions is very similar to the two-dimensional case discussed above and we take the same steps to obtain a three-dimensional version of $\beta_\alpha(\hat{s})$. We start from the 3-D Jacobi–Anger expansion (see [12] page 31)

$$e^{i\alpha\hat{s}\cdot\hat{d}} = \sum_{\ell=0}^{\infty} i^\ell(2\ell+1)j_\ell(\alpha)P_\ell(\hat{d}\cdot\hat{s}), \tag{25}$$

or

$$e^{i\alpha\hat{s}\cdot\hat{d}} = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} 4\pi i^\ell j_\ell(\alpha)Y_{\ell,m}(\hat{s})\overline{Y_{\ell,m}}(\hat{d}). \tag{26}$$

In those expressions, several special functions appear:

- $j_\ell(v)$, the spherical Bessel function of order ℓ. It is linked to the Bessel function via

$$(27) \qquad j_\ell(t) = \sqrt{\frac{\pi}{2t}}\, J_{\ell+\frac{1}{2}}(t).$$

- $P_\ell(x)$, the Legendre polynomial of order ℓ:

$$\begin{cases} P_0(x) = 1, \; P_1(x) = x, \\ (\ell+1)P_{\ell+1}(x) - (2\ell+1)xP_\ell(x) + \ell P_{\ell-1}(x) = 0. \end{cases}$$

Two important properties of these polynomials are [25], [26]

$$(28) \qquad P_\ell(x) \leq 1, \quad \forall x \in [-1,1],$$

$$(29) \qquad P_\ell(x) \leq \frac{2}{\sqrt{\pi(2\ell+1)}} \frac{1}{(1-x^2)^{\frac{1}{4}}}, \quad \forall x \in]-1,1[.$$

- $Y_\ell^m(\hat{s})$, the spherical harmonics of non negative index ℓ and of momentum m, where m varies from $-\ell$ to ℓ; the set $\{Y_{\ell,m}(\hat{s})\}_{|m|\leq\ell<\infty}$ forms a complete orthonormal system in $L^2(S^2)$. Let θ, φ be the spherical angles defined for $\hat{s} = (s_1, s_2, s_3)^T \in S^2$ by the expressions

$$\hat{s}_1(\theta,\varphi) = \cos\varphi \sin\theta, \quad \hat{s}_2(\theta,\varphi) = \sin\varphi\sin\theta, \quad \hat{s}_3(\theta,\varphi) = \cos\theta.$$

In these coordinates the spherical harmonics are given by

$$(30) \qquad Y_\ell^m(\theta,\varphi) = \sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-|m|)!}{(\ell+|m|)!}}\, P_\ell^{|m|}(\cos\theta)\, e^{im\varphi},$$

where $P_\ell^m(t)$ are the associated Legendre functions defined by

$$P_\ell^m(x) = (1-x^2)^{\frac{m}{2}} \frac{d^m P_\ell(x)}{dx^m}, \qquad m \geq 0.$$

The equivalence between (25) and (26) comes from the additional formula

$$(31) \qquad (2\ell+1)P_\ell(\hat{d}\cdot\hat{s}) = 4\pi \sum_{m=-\ell}^{\ell} Y_\ell^m(\hat{s})\overline{Y_\ell^m}(\hat{d}).$$

Inserting the Jacobi–Anger expansion (26) into (12), we get

$$U_\alpha^{\text{ray}}(\hat{s}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} 4\pi i^\ell j_\ell(\alpha) \left(\sum_{n=1}^{N} B_n \overline{Y_\ell^m}(\hat{d_n}) \right) Y_\ell^m(\hat{s})$$

As in 2-D, it can be shown that if α is not too large, say $\alpha < 100$, the Jacobi–Anger series can be truncated at $|\ell|$ less than $L(\alpha)$, see (16). The constants C_1 and C_2 can be chosen approximatively as in the 2-D case, although the truncation error in the Jacobi Anger series increases sligtly when going from 2-D to 3-D, [9].

Next, we introduce the space of spherical Fourier coefficients

$$l^2_{\text{sphere}} = \left\{ \gamma_{\ell,m},\, 0 \le \ell < \infty,\, -\ell \le m \le +\ell,\, \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} |\gamma_{\ell,m}|^2 < \infty \right\},$$

and let $\mathcal{F}_{\text{sphere}} : L^2(S^2) \to l^2_{\text{sphere}}$ be the mapping

$$(\mathcal{F}_{\text{sphere}} A)_{\ell,m} := \int_{S^2} A(\hat{s}) \overline{Y_\ell^m}(\hat{s}) d\sigma(\hat{s}). \tag{32}$$

with inverse

$$\left(\mathcal{F}^{-1}_{\text{sphere}} \gamma\right)(\hat{s}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \gamma_{\ell,m} Y_\ell^m(\hat{s}).$$

Moreover, assuming that $j_\ell(\alpha) \neq 0$ for $|\ell| \le L(\alpha)$, we introduce the dividing operator $D_\alpha : l^2_{\text{sphere}} \to l^2_{\text{sphere}}$, given by

$$(D_\alpha \gamma)_{\ell,m} := d_\ell^\alpha \, \gamma_{\ell,m}, \qquad d_\ell^\alpha = \begin{cases} \frac{4\pi}{(L(\alpha)+1)^2} \frac{1}{4\pi i^\ell j_\ell(\alpha)}, & \ell \le L(\alpha), \\ 0, & \text{otherwise.} \end{cases}$$

We then define the three-dimensional β_α-function

$$\beta_\alpha = \mathcal{F}^{-1}_{\text{sphere}} D_\alpha \mathcal{F}_{\text{sphere}} U_\alpha.$$

Hence,

$$\beta_\alpha(\hat{s}) = \sum_{\ell=0}^{L(\alpha)} \sum_{m=-\ell}^{\ell} d_\ell^\alpha \left(\mathcal{F}_{\text{sphere}} U_\alpha\right)_{\ell,m} Y_\ell^m(\hat{s}). \tag{33}$$

For the ray solution (12) we get

$$\begin{aligned}\beta_\alpha^{\text{ray}}(\hat{s}) &= \left(\mathcal{F}_{\text{sphere}}^{-1} D_\alpha \mathcal{F}_{\text{sphere}} U_\alpha^{\text{ray}}\right)(\hat{s}) \\ &= \sum_{n=0}^{N} B_n \frac{4\pi}{(L(\alpha)+1)^2} \sum_{\ell=0}^{L(\alpha)} \sum_{m=-\ell}^{\ell} \overline{Y_\ell^m}(\hat{d}_n) Y_\ell^m(\hat{s}).\end{aligned}$$

Using (31) we get a convenient expression for $\beta_\alpha^{\text{ray}}$ of the same form as in two dimensions,

$$\beta_\alpha^{\text{ray}}(\hat{s}) = \sum_{n=0}^{N} B_n S_\alpha(\hat{s}\cdot\hat{d}_n), \qquad S_\alpha(r) = \sum_{\ell=0}^{L(\alpha)} \frac{2\ell+1}{(L(\alpha)+1)^2} P_\ell(r).$$

To analyze this expression for $\beta_\alpha^{\text{ray}}$, let us first assume that there is only one ray, i.e. $N = 1$. Since $P_\ell(1) = 1$, for all ℓ, we get $S_\alpha(1) = 1$ and consequently,

$$\beta_\alpha^{\text{ray}}(\hat{d}_1) = B_1.$$

When $-1 \leq r < 1$, we can use (29) to get the estimate

$$|S_\alpha(r)| \leq \frac{2}{\sqrt{\pi}\,(1-r^2)^{\frac{1}{4}}\,(L(\alpha)+1)^2} \sum_{\ell=0}^{L(\alpha)} \sqrt{2\ell+1},$$

which implies that there exists a pure constant C such that

$$|S_\alpha(r)| \leq \frac{C}{(1-r^2)^{\frac{1}{4}}\,\sqrt{L(\alpha)+1}}, \qquad r \in]-1,1[. \tag{34}$$

Therefore, $|\beta_\alpha^{\text{ray}}(\hat{s})|$ goes to zero when $\hat{s} \neq \hat{d}_n$ and α goes to infinity. As in two dimensions, we can extend the analysis to multiple rays. Let

$$\text{Sep}(\hat{s}, n) = \frac{\sqrt{1-\left(\hat{s}\cdot\hat{d}_n\right)^2}}{|B_n|^2}.$$

This measures how well a ray in direction $\hat{s}$ is separated from ray n, weighted by the amplitude. Inserted in (34) we have

$$|S_\alpha(\hat{s}\cdot\hat{d}_n)| \leq \frac{C}{|B_n|\sqrt{\text{Sep}(\hat{s},n)[L(\alpha)+1]}}, \qquad \hat{s} \neq \hat{d}_n.$$

The estimates in three dimensions that correspond to (21) and (22) now follows easily,

$$|\beta_\alpha^{\mathrm{ray}}(\hat{d}_n) - B_n| \le \frac{N-1}{\sqrt{L(\alpha)+1}} \max_{m\neq n} \frac{1}{\sqrt{\mathrm{Sep}(\hat{d}_n, m)}}.$$

and

$$|\beta_\alpha^{\mathrm{ray}}(\hat{s})| \le \frac{N}{\sqrt{L(\alpha)+1}} \max_{m} \frac{1}{\sqrt{\mathrm{Sep}(\hat{s}, m)}} \qquad \hat{s} \neq \hat{d}_n,\ \forall n.$$

Hence for α large enough, $\beta_\alpha^{\mathrm{ray}}(\hat{d}_n) \simeq B_n$ and $\beta_\alpha^{\mathrm{ray}}(\hat{s}) \simeq 0$ when $\hat{s} \neq \hat{d}_n$. The limit as α goes to infinity is the same as the two-dimensional case given in (23). The analysis hence confirms that the three-dimensional β_α defined on the sphere is also a function with sharp peaks in the ray directions when α is large enough.

3 Algorithm and numerical results in 2-D

The numerical algorithm is broken up into several steps. The key to finding the unknown number of rays, their complex amplitudes and directions is the function β_α introduced in Section 2. As was seen there, this function will be close to zero away from the dominant ray directions, and have a value close to the complex amplitude B_n in the direction of ray n. In our algorithm, we therefore first compute a numerical approximation of $\beta_\alpha(\hat{s})$ on a uniform discretization of the unit circle, possibly including a regularization procedure. Second, we analyze the numerical results to find a preliminary set of ray directions and amplitudes. Finally, we post-process this set via a nonlinear optimization procedure to get a more accurate result.

3.1 Approximation of β_α

We want to approximate the function $\beta_\alpha(\hat{s})$ numerically in order to determine the directions and amplitudes of the rays. As was discussed in Section 2, this function should have strong maxima in the directions of the rays. Moreover, the value of β_α itself in these directions approximates the corresponding complex amplitudes. The algorithm is straightforward, using the fast Fourier transform :

Introduce a uniform grid $\{\tilde{\theta}_m\}$ with $M \geq 2L(\alpha)+1$ points on the unit circle,

$$\tilde{\theta}_m = m\Delta\theta, \qquad \Delta\theta = \frac{2\pi}{M}, \qquad m = 0, \ldots, M-1. \tag{35}$$

Then let $\{\tilde{U}_m\}$ be the grid function that samples the given function $U_\alpha(\hat{s})$ in the grid points,

$$\tilde{U}_m = U_\alpha(\tilde{d}_m), \qquad \tilde{d}_m = (\cos\tilde{\theta}_m, \sin\tilde{\theta}_m). \tag{36}$$

Assuming U_α is well-approximated by its Fourier interpolant in those points we can then compute an approximation of β_α as follows:

$$\{\hat{U}_\ell\} = \mathrm{FFT}\left\{\tilde{U}_m\right\} \qquad \{b_m\} = \mathrm{FFT}^{-1}\left\{d_\ell^\alpha \hat{U}_\ell\right\}, \qquad \beta_\alpha(\tilde{\theta}_m) \simeq b_m$$

where d_ℓ^α was defined in (17). This is the discrete version of the $\mathcal{F}^{-1}D_\alpha\mathcal{F}$ operator of Section 2. Since the frequencies $\hat{U}_\ell$ are multiplied by the divisors d_ℓ^α which vanish for $\ell \notin [-L(\alpha), L(\alpha)]$, it is also clear that there is no point in taking $M > 2L(\alpha)+1$. Therefore, we take precisely

$$M = 2L(\alpha)+1, \tag{37}$$

and finally get

$$\{b_m\} = 2\pi\mathrm{FFT}^{-1}\left\{\frac{\mathrm{FFT}\left\{\tilde{U}_m\right\}}{(2L(\alpha)+1)i^\ell J_\ell(\alpha)}\right\}. \tag{38}$$

We end up with only one free parameter, α, which specifies both the radius of the observation circle around x_0 and the truncation and discretization parameters according to (16) and (37).

Remark 2. *Going back to the original assumptions (11) and (12), we see that with the same discretization as above, they reduce to the linear system of equations*

$$\tilde{U}_\ell = \sum_{m=0}^{M-1} \tilde{b}_m e^{i\alpha\, \tilde{d}_\ell \cdot \tilde{d}_m}, \qquad \ell = 0, \ldots, M-1.$$

The system matrix is a circulant matrix and as an alternative we could solve it directly at a comparable cost, $\mathcal{O}(M\log M)$ [19]. However, the condition number of this matrix rapidly deteriorates when M grows, and also in this case one finds that the problem becomes very ill-conditioned if $M > 2L(\alpha)+1$. The FFT-based Jacobi–Anger inversion is in fact a stabilized, approximation ($b_m \approx \tilde{b}_m$) of the standard fast way to solve this circulant matrix problem.

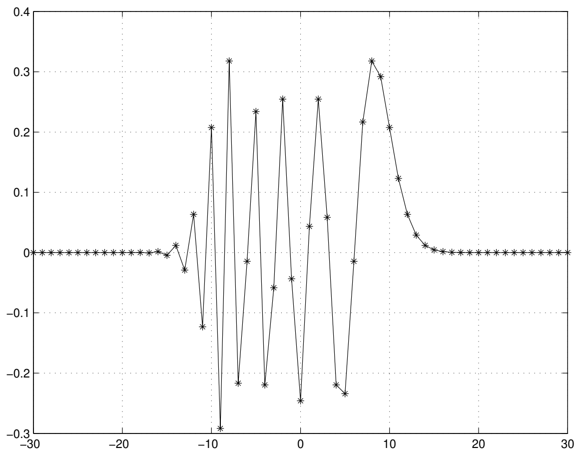


Figure 2: The Bessel function $J_\ell(\alpha)$ as a function of ℓ for fixed $\alpha = 10$. Stars indicates values for integers ℓ.

3.2 Tichonov regularisation

The previous subsection gives a rule how to safely truncate the Jacobi–Anger series. We can then get rid of the small values of the Bessel functions when $\ell \to \pm\infty$. This is very important as the step (38) of the algorithm involves a division by $J_\ell(\alpha)$ and we indeed assumed on this occasion that $J_\ell(\alpha) \neq 0$. This last condition can be violated as Bessel functions as functions of α for instance have zeros and we cannot guarantee that α is not close to or exactly one of them, c.f. Figure 2. If the truncation is too large the exponential decay can also lead to very small coefficients that impair the precision of formula (38).

A simple solution to this problem is to use a Tichonov type regularization. Let us write the last step in the FFT algorithm, $\{b_m\} = \mathrm{FFT}^{-1}\{\hat{U}_\ell d_\ell^\alpha\}$, as a linear system with $b = \{b_m\}$ as the unknown,

$$Gb = \hat{U}, \qquad G = \{g_{\ell m}\}, \qquad g_{\ell m} = (2L(\alpha)+1) i^\ell J_\ell(\alpha) e^{-i\ell\tilde{\theta}_m}, \qquad \hat{U} = \{\hat{U}_\ell\}.$$

It is clear that zeros of the Bessel function J_ℓ may be a problem for the numerical resolution of this system. So, instead we propose to solve the "regularized" system

$$(G^*G + \varepsilon I) b_\varepsilon = G^* \hat{U},$$

where I is the identity matrix. The inversion formula (38) becomes

$$\{b_m^\varepsilon\} = 2\pi \mathrm{FFT}^{-1}\left(\left\{\frac{\hat{U}_\ell(2L(\alpha)+1)J_\ell(\alpha)}{i^\ell[(2L(\alpha)+1)^2 J_\ell(\alpha)^2 + 4\varepsilon\pi^2]}\right\}\right). \tag{39}$$

and remains correct even when $J_\ell(\alpha)$ is zero or close to zero.

The exact Tichonov regularization would consist in choosing ε such that the relative error between the actual and regularized problem is smaller than a prescribed precision (an optimization problem must then be solved).

3.3 Preliminary processing and accuracy

Let $b_{\min}$ be a given tolerance parameter. A simple way to determine a relevant number of ray directions from our approximation of β_α is to check directions $\tilde{\theta}_m$ for which the corresponding b_m coefficient satisfies

$$|b_m| > |b_{m-1}|, \qquad |b_m| > |b_{m+1}|, \qquad |b_m| > b_{\min},$$

Hence, we select all local maxima in m whose amplitudes are sufficiently large. Suppose that we get $\tilde{N} \simeq N$ directions, $\tilde{\theta}_m$, with $m = m_1, \dots, m_{\tilde{N}}$. Close to θ_n we know that $\beta_\alpha(\hat{s}(\theta)) \simeq B_n S_\alpha(\theta - \theta_n)$ by (21) in Section 2. Therefore, assuming that $\theta_n = \tilde{\theta}_{m_n} + \eta\Delta\theta$ for some $\eta \in (-1, 1)$,

$$b_{m_n} \simeq B_n S_\alpha(\tilde{\theta}_{m_n} - \theta_n) = B_n \frac{\sin(\eta\pi)}{M\sin\left(\frac{\eta\pi}{M}\right)} \simeq B_n \frac{\sin(\eta\pi)}{\eta\pi}. \tag{40}$$

We then get

$$\frac{b_{m_n}}{b_{m_n+1}} \simeq \frac{(\eta-1)\pi\sin(\eta\pi)}{\eta\pi\sin((\eta-1)\pi)} = \frac{1-\eta}{\eta},$$

and consequently,

$$\eta \simeq \frac{b_{m_n+1}}{b_{m_n} + b_{m_n+1}}.$$

We hence compute the preliminary ray directions and complex amplitudes as

$$\tilde{\eta} = \Re\left[\frac{b_{m_n+1}}{b_{m_n} + b_{m_n+1}}\right], \qquad \theta_n^{\mathrm{prel}} = \tilde{\theta}_{m_n} + \tilde{\eta}\Delta\theta, \qquad B_n^{\mathrm{prel}} = \frac{b_{m_n}\tilde{\eta}\pi}{\sin(\tilde{\eta}\pi)}.$$

and conclude that, close to $x = x_0$,

$$u_k(x) \simeq \sum_{n=1}^{\tilde{N}} B_n^{\mathrm{prel}} e^{ik\eta(x_0)(x-x_0)\cdot\hat{d}_n^{\mathrm{prel}}}, \qquad \hat{d}_n^{\mathrm{prel}} = (\cos\theta_n^{\mathrm{prel}}, \sin\theta_n^{\mathrm{prel}}).$$

We can thus efficiently compute an approximate set of ray directions and complex amplitudes through the FFT-based algorithm and the simple selection algorithm above. There are a number of error sources in this computation that we need to be aware of:

1. High frequency approximation in (8).

2. Linearization in (10).

3. Approximation of $U_\alpha(\hat{s})$ using M samples in (36).

4. Approximation of β_α locally by a sinc function in (40).

The contribution from the first error source decreases with the frequency. The second error source increases with α and decreases with frequency. By the scaling of U_α, the effect of the remaining sources are *independent of the frequency.* Error source three obviously decreases with increasing M. The last source decreases with increasing α and increases when the rays are not well separated by (21). In total, we thus have a method for which the accuracy improves when the frequency increases, and the amount of work is held fixed (fixed α, M).

For reasonably high frequencies the last two sources dominate, and since $M = 2L(\alpha) + 1$ by (37), only the free parameter α determines the accuracy. In general the larger the observation circle, the finer the discretization and the better the accuracy of our result. This is the limitation in the localization of our procedure: there is no hope to recover the direction of the rays at a single point without information in a neighborhood which allows for some correlations. If, however, the neighborhood exceeds the domain of validity of the local plane wave approximation (4) then the "ansatz" assumption (6) may not be relevant anymore. Note that for a sum of plane waves *only* the last two sources contribute to the error.

3.4 Post-processing

As already discussed, the precision of our method is limited by the size of the "observation" circle around x_0 and this may be a severe restriction. So we propose here a post-processing procedure where the data obtained from the spectral inversion is used as initial data for a routine that tries to fit the expansion (12) directly to the sampled solution values. This is done by nonlinear minimization of the residual. The accuracy of the rays' complex amplitudes and directions can then be significantly improved.

We define the residuals

$$r_m = \tilde{U}_m - \sum_{n=1}^{\tilde{N}} \bar{B}_n e^{i\alpha \tilde{d}_m \cdot \bar{d}_n}, \qquad m = 0, \dots, M-1,$$

where $r = (r_0, \dots, r_{M-1})^T \in \mathbb{C}^M$ depends on the parameters $\bar{B}_n$ and $\bar{d}_n$. For the correct choice of those parameters, the residual should be small by our assumptions. We thus try to minimize the norm of r by varying the parameters,

$$\min_{\bar{B}_n, \bar{d}_n,} ||r||.$$

We solve this nonlinear problem with the standard Gauss–Newton minimization algorithm, using the preliminary values as starting values

$$\bar{B}_n^0 = B_n^{\text{prel}}, \qquad \bar{d}_n^0 = \hat{d}_n^{\text{prel}}, \qquad n = 1, \dots, \tilde{N}.$$

Usually a few iterations gives a dramatic improvement in the accuracy of the results (see Section 3.5.2). We denote the post-processed result B_n^{post} and θ_n^{post}.

Note that this optimization procedure requires good starting values to converge. The preliminary results must therefore be fairly accurate. By the discussion in the previous section, this means that we need to take a sufficiently large α (i.e. M) when we compute the preliminary results, and that it must be larger when the rays are not well separated. Also note that the magnitude of the final residual after minimization is a measure of the quality of the GO approximation at the frequency k.

In the post-processing one could also use other points than those on the observation circle, if they are readily available. One could possibly also include higher order terms in the linearization (10) to reduce the error from source two, in the Section 3.3 discussion.

3.5 Numerical results in 2-D

3.5.1 Point sources solutions in homogeneous space $\eta = 1$

We consider the following source points problem

$$\Delta u_k + k^2 u_k = \sum_{n=1}^{N} 4i\sqrt{k}\, a_n \delta_{x_n}. \tag{41}$$

The dirac masses are centered at the source points (x_n), modulated by the amplitudes (a_n) and normalized such that the solution is given as the sum of Hankel functions with decaying amplitudes (asymptotically independent of k), centered at points x_n :

$$u_k(x) = \sum_{n=1}^{N} a_n \sqrt{k} H_0^1(k|x - x_n|).$$

We want to recover ray directions and complex amplitudes at the observation point $(0,0)$. Let us specify the source points in polar coordinates $x_n = (R_n \cos\bar{\theta}_n, R_n \sin\bar{\theta}_n)$. For large $t = kr$ we have

$$H_0^1(t) \simeq e^{i\left(t-\frac{\pi}{4}\right)}\sqrt{\frac{2}{t\pi}},$$

and we therefore take

$$B_n = a_n \sqrt{\frac{2}{R_n \pi}} \exp\left(i\left(kR_n - \frac{\pi}{4}\right)\right), \qquad \theta_n = \bar{\theta}_n + \pi,$$

as the "exact", asymptotic complex amplitude value and ray directions. The wavenumber in these computations was $k = 10^4$.

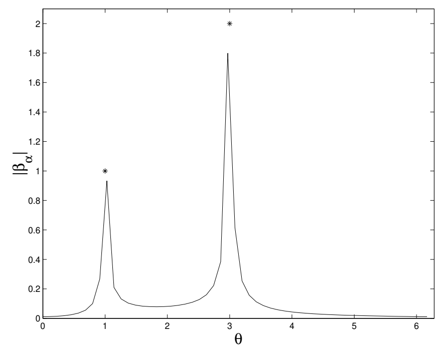
In Figure 3 we show the result obtained for different number of source points. The computed amplitudes (vertical axis) are plotted as a function of the angle (horizontal axis). The stars indicate the exact angles and asymptotic amplitudes $(\theta_n, |B_n|)$ at the observation point. We used $\alpha/2\pi = 1.5$. Upon simplifying (16) into $L(\alpha) = \alpha + 8\log(\alpha)$ this corresponds to $M = 55$.

The case with five sources is further illustrated in Figure 4, where the real part of the actual solution is plotted together with the observation circle and a polar plot of $|\beta_\alpha|$.

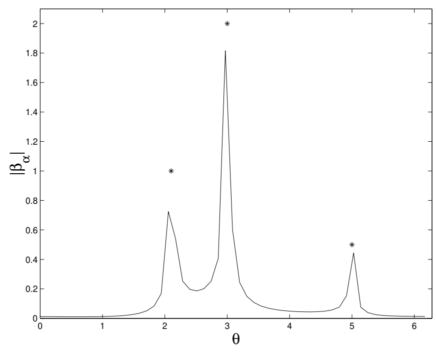
Figure 5 shows the nine source case when computed with variable size α of the observation circle. It is clear that the precision of the predicted ray directions increases with α, as expected.

3.5.2 Post-processing and convergence analysis

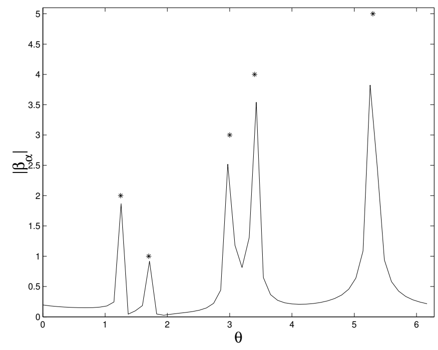
One drawback of the basic method is the constraint which links the angle discretization with the size of the circle on which we sample the solution. This circle should not be too large, since the first order Taylor expansion of the phase and amplitude around the observation point deteriorates when the radius of the observation circle gets large. Moreover, we may have to retrieve GO components near scattering bodies. Hence the idea (explained) in Section 3.4 to apply the algorithm on a discretization of angles that is



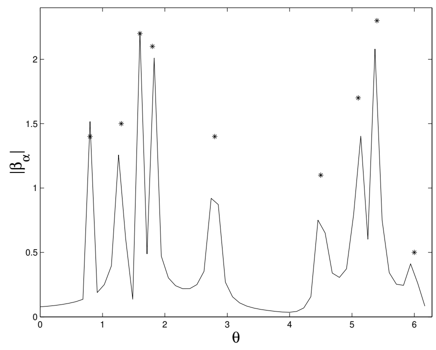
(a) 2 sources



(b) 3 sources

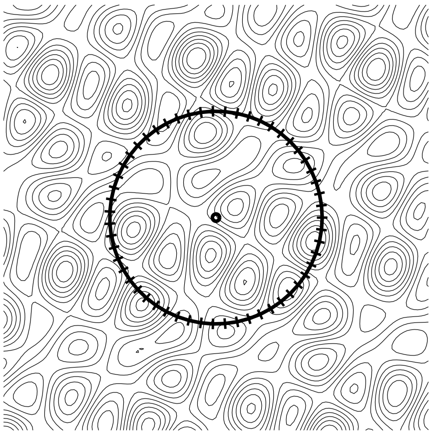


(c) 5 sources

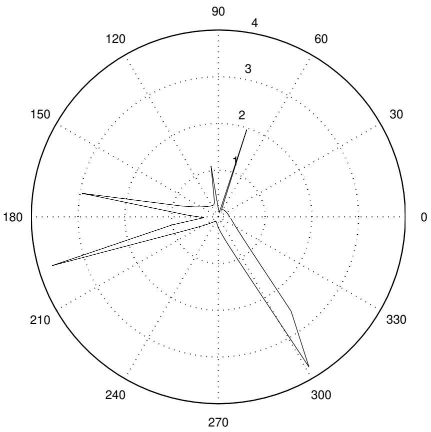


(d) 9 sources

Figure 3: Numerical approximation of $|\beta_\alpha(\hat{s}(\theta))|$ with $k = 10^4$, $\alpha/2\pi = 1.5$ and $M = 55$ and variable number of sources N. Stars indicate the exact asymptotic values of $|B_n|$ and θ_n when $k \to \infty$.

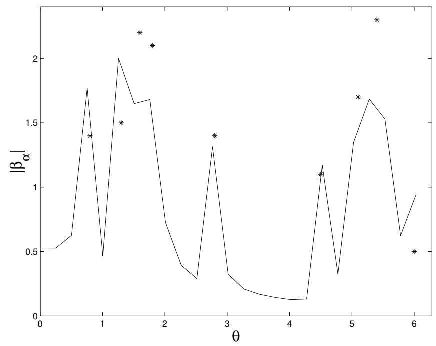


(a) Full solution and observation circle.

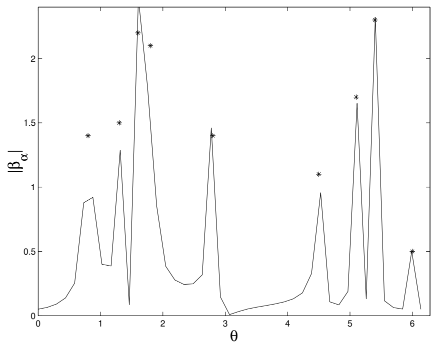


(b) Approximation of $|\beta_\alpha(\hat{s}(\theta))|$.

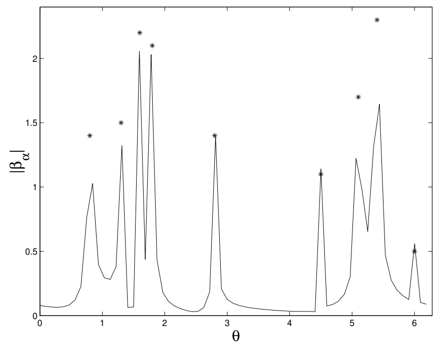
Figure 4: Example in Figure 3 with five sources, $\alpha/2\pi = 1.5$, $M = 55$ and $k = 10^4$. Left figure shows a contour plot of the real part of the solution $u_k(x)$ with the observation circle $|x - x_0| = \alpha/k$ and its discretization superimposed. Right figure shows a polar plot of the numerical approximation of $|\beta_\alpha(\hat{s}(\theta))|$.



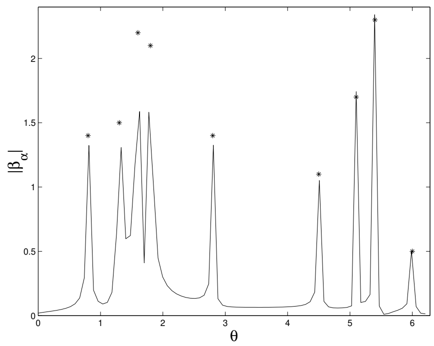
(a) $\alpha/2\pi = 0.5,\ M = 25$



(b) $\alpha/2\pi = 1,\ M = 43$



(c) $\alpha/2\pi = 2,\ M = 67$



(d) $\alpha/2\pi = 3,\ M = 85$

Figure 5: Numerical approximation of $|\beta_\alpha(\hat{s}(\theta))|$ with nine sources, $k = 10^4$ and variable α and M. Stars indicate the exact asymptotic values of $|B_n|$ and θ_n when $k \to \infty$.

coarse, but still sufficient to identify the number of rays N and an approximate set of ray directions and complex amplitudes. Then this approximate solution is used as the initial data for a non-linear inversion method.

A typical improvement in the results after post-processing is given in Table 1 where the results of the five source case is given in more detail, before and after post-processing. The error is reduced by a factor of hundred for this problem.

In Figure 6, a more systematic convergence study is made for the same problem. There, the maximum error in the ray directions, θ_n, the complex amplitudes, B_n and the modulus of the amplitudes $A_n = |B_n|$ are plotted as a function of α ($\sim$ the observation circle radius) as well as of k ($\sim$ the frequency). Both the error in the preliminary results and the post-processed results are shown.

In Figure 6a one can see that the preliminary results improve with α. The post-processed results are much more accurate, but worsen with α. Also note that the accuracy of A_n is much better than that of B_n. Figure 6b shows that the preliminary results are essentially independent of frequency, while the post-processed results improve markedly for higher frequencies.

Our interpretation of the results is that for the preliminary directions and amplitudes, the last two error sources discussed in Section 3.3 dominate. Post-processing eliminates these error sources, and the error after post-processing only comes from the first two sources.

3.5.3 Scattering by a hard disk

We apply our complete procedure (Jacobi–Anger inversion + non linear post-processing) to the scattering of a plane wave by a hard disk of radius $a = 1$ and center at the origin. In this case the input is the exact scattered solution u_k given by the formula ([6] p. 376),

$$u_k(x) = e^{ikr\cos\theta} - \sum_{l=-\infty}^{+\infty} i^k \frac{J_\ell(ka)}{H_\ell^{(1)}(ka)} e^{il\theta} H_\ell^{(1)}(k\,r),$$

where $x = r(\cos\theta, \sin\theta)$ and $J_\ell(t)$, $H_\ell^{(1)}(t)$ are respectively the Bessel and Hankel function of the first kind. (In practice, this series is truncated to the ℓ which modulus are less than some number larger than k.)

The asymptotic solution can be described in terms of rays [7] :

- The incident rays associated to the plane wave arriving from the left.

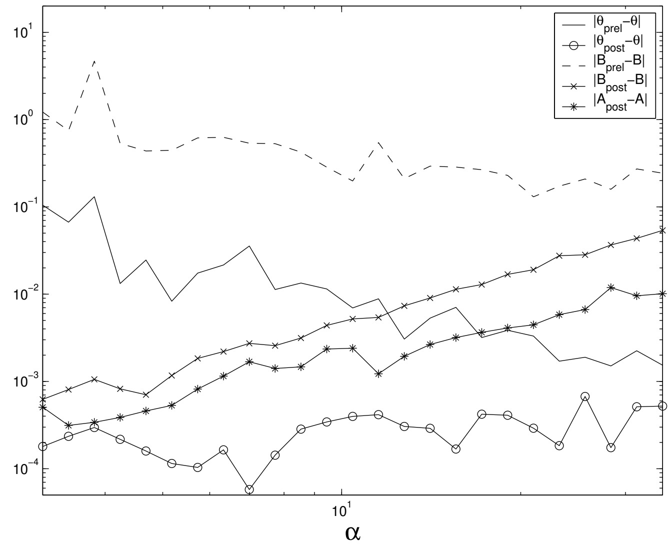
Ray directions

n	Exact θ_n	Preliminary θ_n^{prel}	$\lvert\theta_n^{\mathrm{prel}} - \theta_n\rvert$	Post-processed θ_n^{post}	$\lvert\theta_n^{\mathrm{post}} - \theta_n\rvert$
1	1.25	1.2530	$0.297 \cdot 10^{-2}$	1.249921	$0.79 \cdot 10^{-4}$
2	1.70	1.7116	$1.158 \cdot 10^{-2}$	1.699656	$3.44 \cdot 10^{-4}$
3	3.00	3.0038	$0.379 \cdot 10^{-2}$	2.999926	$0.74 \cdot 10^{-4}$
4	3.40	3.4008	$0.079 \cdot 10^{-2}$	3.399936	$0.64 \cdot 10^{-4}$
5	5.30	5.2990	$0.096 \cdot 10^{-2}$	5.299990	$0.10 \cdot 10^{-4}$

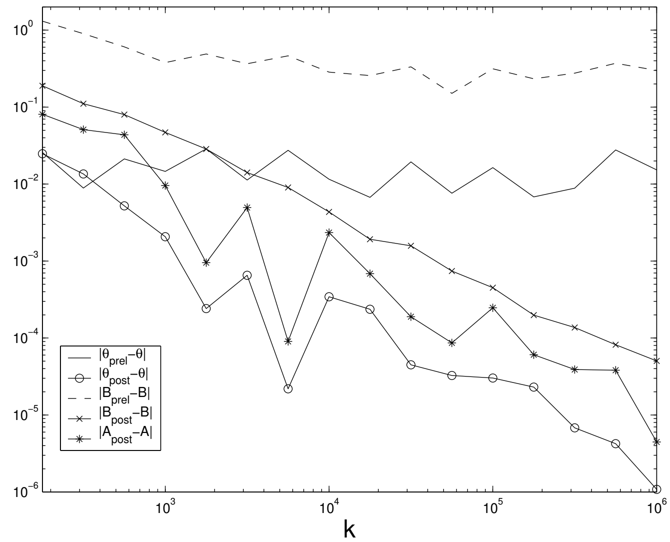
Complex amplitudes

n	Exact B_n	Preliminary B_n^{prel}	$\lvert B_n^{\mathrm{prel}} - B_n\rvert$	Post-processed B_n^{post}	$\lvert B_n^{\mathrm{post}} - B_n\rvert$
1	$-1.7788 + 0.9143i$	$-1.7794 + 0.7861i$	0.1283	$-1.7796 + 0.9103i$	$4.15 \cdot 10^{-3}$
2	$0.3141 + 0.9494i$	$0.2637 + 0.8550i$	0.1070	$0.3141 + 0.9514i$	$1.99 \cdot 10^{-3}$
3	$2.9596 - 0.4905i$	$2.9851 - 0.2068i$	0.2848	$2.9581 - 0.4887i$	$2.34 \cdot 10^{-3}$
4	$-0.0358 - 3.9998i$	$-0.3082 - 3.9398i$	0.2790	$-0.0332 - 4.0022i$	$3.47 \cdot 10^{-3}$
5	$-4.9465 - 0.7292i$	$-4.9423 - 0.8172i$	0.0882	$-4.9465 - 0.7335i$	$4.35 \cdot 10^{-3}$

Table 1: Test problem with five sources, $k = 10^4$, $\alpha/2\pi = 1.5$, $M = 55$.

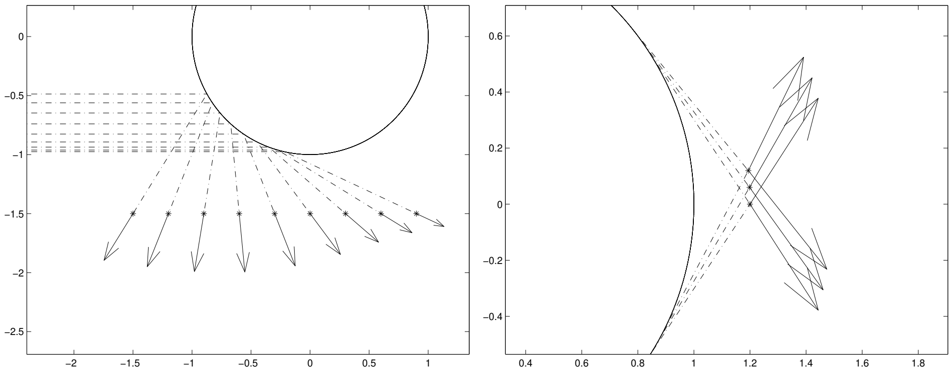


(a) Variable α, fixed $k = 10^4$.



(b) Variable k, fixed $\alpha/2\pi = 1.5$.

Figure 6: Convergence. Five sources as in table. $b_{\min} = 0.2$



(a) Reflected rays.

(b) Diffracted rays.

Figure 7: Reflection and diffraction by a disk.

- Rays reflected symmetrically with respect to the normal to the disk (see Figure 7a).

- Diffracted rays, which after reaching the disk tangentially, creep on its surface and then initiate tangent rays at all subsequent points of its boundary (see Figure 7b).

We used our numerical method to compute these reflected and diffracted rays, with $k = 100$, $\alpha = 5$ and $M = 37$. Figure 7 shows the results at a collection of points below and just behind the disk. The arrows indicate the numerically computed directions of the rays and the length of the arrows are scaled according the computed amplitudes. The dashed lines follow the exact direction of the reflected and the diffracted rays.

We now select point $(1.2, 0.15)$ where diffraction can be observed and try to estimate the frequency dependence of the amplitude numerically. We indeed know [23] that at such points the amplitude of a diffracted ray behaves as

$$A(k) = C_1 k^{-\frac{1}{6}} e^{-C_2 (\frac{k}{2})^{\frac{1}{3}}},$$

where C_1 and C_2 are frequency independent constants. So we select a ray and compute the amplitudes for values of k going from 100 to 1000, using

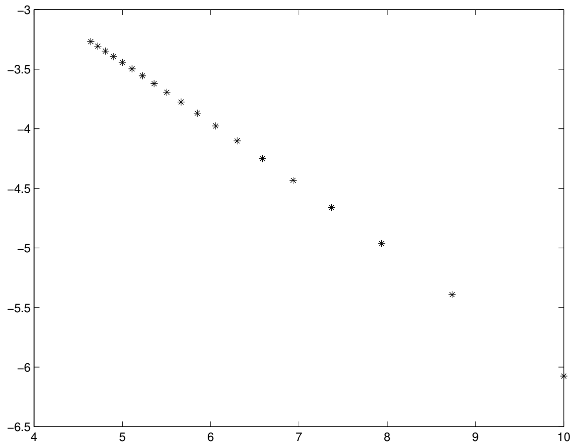


Figure 8: Frequency dependence of the diffracted amplitude. Horizontal axis : $k^{\frac{1}{3}}$. Vertical axis : $\log(|A(k)|k^{\frac{1}{6}})$, where $A(k)$ has been computed by the algorithm.

$\alpha = 10$ for all computations. In Figure 8 we plot the function

$$k^{\frac{1}{3}} \mapsto \log(|A|k^{\frac{1}{6}}) = \log C_1 - C_2\,(\frac{k}{2})^{\frac{1}{3}},$$

and as expected, a linear behavior is obtained.

3.5.4 Fold caustic solution in heterogeneous medium

The algorithm is local and also works for inhomogeneous media where the index of refraction η is not constant. As we need to compute an Helmholtz solution for possibly large values of k we choose an example where the index only depends on one of the Cartesian coordinate $x = (x_1, x_2)$, namely

$$\eta(x) = \begin{cases} \frac{1}{2}\,(\cos(\pi x_1) + 1), & x_1 \in]0, 2[, \\ 1, & \text{otherwise.} \end{cases}$$

One can then apply the method of separation of variables. We search for a solution that is an incident plane wave $\exp(ik\cos\alpha x_1 + \sin\alpha x_2))$ plus a diffracted wave of the form $u_d(x) = \tilde{u}(x_1)\exp(ikx_2 \sin\alpha)$ where $\tilde{u}$ satisfies the 1-D Helmholtz equation

$$\frac{d^2\tilde{u}}{dx_1^2}(x_1) + k^2(\eta^2(x_1) - \sin^2\alpha)\tilde{u}(x_1) = k^2(1 - \eta^2(x_1)). \tag{42}$$

Away from the zone of varying index $x_1 \in]0,2[$ we can derive exact transparent boundary conditions for $\tilde{u}$ and discretize (42) using a finite difference method on a bounded domain. Even using direct inversion of the matrix makes it possible to take k as large as 100 for 10 discretization points per wavelength.

The initialization process therefore consists in computing the Helmholtz solution on a 2-D finite difference grid as explained above. Then the input to our numerical method simply is a linear 2-D interpolation between those points to estimate the solution on a circle around the observation point. We show in Figure 9 the output (using $\alpha = 10$) at a collection of points approaching the caustic, superimposed on the real part of the Helmholtz solution. The arrows again indicate the numerically computed directions of the rays and the length of the arrows are scaled according the amplitudes. In this situation (see [5]) geometrical optics predicts two rays with phases satisfying the Eikonal equations

$$\phi^{\pm}_{x_2} = \pm\sqrt{\eta^2 - \phi^{\pm}_{x_1}}. \tag{43}$$

Based on (43), the directions of the rays are easily found and are represented by the dashed lines in Figure 9. The results also show an increase of the amplitude as we approach the caustic which is consistent with the GO model.

4 Algorithm and Numerical results in 3-D

In this section we indicate how the 2-D algorithm can be generalized to the 3-D case.

4.1 Approximation of β_α

The approximation of the three-dimensional β_α in (33) is done in essentially the same way is in two dimensions. The main difference is that for practical reasons we cannot compute the spherical Fourier coefficients exactly. Instead we use a quadrature rule on the sphere to evaluate the integral in (32) when we compute the coefficients.

The quadrature rule is chosen such that the spherical harmonics of order less than $2L(\alpha)$ are integrated exactly. There are several possible ways to achieve that. Considering that $d\sigma(\hat{s}) = \sin\theta d\theta d\varphi = d\cos\theta d\varphi = dy d\varphi$, we use Gauss-Legendre quadrature in the y variable and the trapezoidal rule in the (periodic) φ variable. The Gauss-Legendre rule has $N_\theta = L(\alpha)+1$ nodes

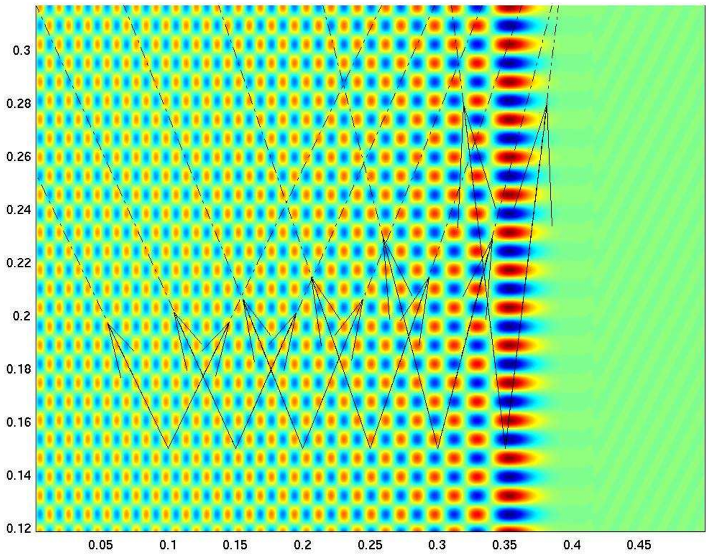


Figure 9: Rays near a caustic

in the points $\{y_i\}_{i=0}^{N_\theta-1} \subset [-1,1]$. The corresponding angles are $\theta_i = \arccos y_i$ and we denote the corresponding weights by ω_θ^i. For the trapezoidal rule, we use $N_\varphi = 2(L(\alpha)+1)$ equidistant nodes $\{\varphi_j\}_{j=0}^{N_\varphi-1} \subset [0,2\pi]$, where $\varphi_j = 2\pi j/N_\varphi$. The corresponding weights are simply $\omega_\varphi^j = 2\pi/N_\varphi$. The quadrature rule then reads

$$\oint_{S^2} F(\hat{s})d\sigma(\hat{s}) := \sum_{i=0}^{N_\theta-1}\sum_{j=0}^{N_\varphi-1} \omega_\varphi^j \omega_\theta^i \, F\left(\hat{s}(\theta_i,\varphi_j)\right).$$

Assuming that we can sample the solution via the function $U_\alpha(\hat{s})$ in the points $\hat{s} = \hat{s}(\theta_i,\varphi_j)$, we can then use the rule to approximate the spherical Fourier coefficients, in (33) as

$$(\mathcal{F}_{\text{sphere}}U_\alpha)_{\ell,m} \simeq \oint_{S^2} U_\alpha(\hat{s})\overline{Y_\ell^m}(\hat{s})d\sigma(\hat{s}).$$

Once these are computed, formula (33) is used to evaluate $\beta_\alpha(\hat{s})$. Like in two dimensions, the procedure may need to be regularized. The processing steps in Section 3.3 and Section 3.4 could also be extended to three dimensions in the obvious way.

4.2 Numerical results in 3-D

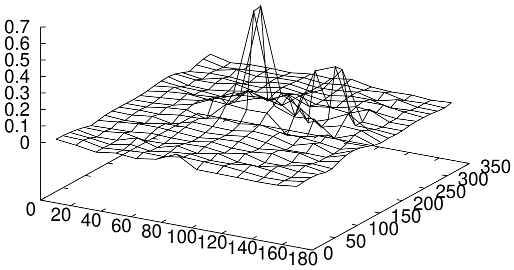
We consider a simple problem: the solution to the Helmholtz equation in a homogeneous domain with two point sources.

$$u_k(x) = a_1 \frac{|x_1^s|\, e^{ik|x-x_1^s|}}{|x-x_1^s|} + a_2 \frac{|x_2^s|\, e^{ik|x-x_2^s|}}{|x-x_2^s|}.$$

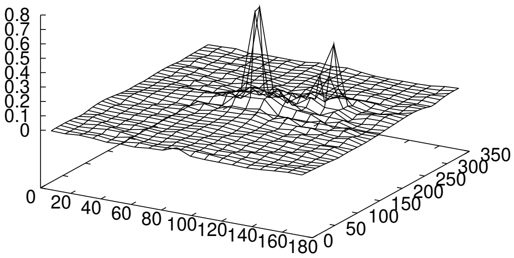
We observe the solution at the origin, $x_0 = 0$. When k is large we obtain two rays with

$$A_n(x_0) = a_n, \quad \hat{d}_n(x_0) = -\frac{x_n^s}{|x_n^s|}, \quad n = 1,2.$$

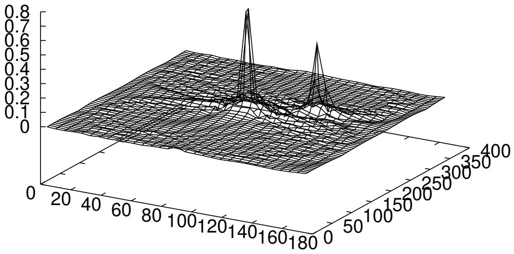
We choose the amplitudes $a_1 = 1$, $a_2 = \frac{1}{2}$, the source locations $x_1^s = (1,0,0)$, $x_2^s = \frac{1}{\sqrt{3}}(1,1,1)$ and the wavenumber $k = 10^4$. The results are displayed in Figure 10. The algorithm seems to be able to reconstruct the angles of propagation, and the resolution improves with α.



(a) $\alpha = 5$



(b) $\alpha = 10$



(c) $\alpha = 20$

Figure 10: Function $|\beta_\alpha(\hat{s})|$ for two point sources of amplitudes 1 and 0.5 and various values of α; the source locations $x_1^s = (1, 0, 0)$, $x_2^s = \frac{1}{\sqrt{3}}(1, 1, 1)$, the observation point $x_0 = (0, 0, 0)$ and the wavenumber $k = 10^4$.

5 Conclusion

We presented an algorithm which given a fixed frequency domain solution in a local neighborhood of an "observation" point computes its GO asymptotic representation around this point. The algorithm is cheap : $\mathcal{O}(M \log M)$ operations is needed, where M is the number of data samples of the Helmholtz solution on a small circle around the observation point. This number depends on the radius of the circle through (37), but it is independent of the frequency. A larger number is necessary to obtain convergence when the separation between rays is small. The accuracy of the method increases with the frequency. It is easy to implement and can be applied both for heterogeneous and homogeneous media. A remarkable feature is its ability to capture diffracted rays.

Integral equation method are well suited for frequency domain calculations in homogeneous media. For scattering problem the solution is given as currents (i.e. functions defined) on the boundary of the scattering object. If the input to our method is given in that form there is a way to reduce the cost of the expensive integral field calculations needed to evaluate the solution locally around the observation point. The amplitude coefficients can indeed be expressed as Herglotz waves (see [12]) which only depend on the currents. This approach will be explained in a forthcoming paper.

References

[1] T. Abboud, J.-C. Nédélec, and B. Zhou. Méthode des équations intégrales pour les hautes fréquences. *C. R. Acad. Sci. Paris Sér. I Math.*, 318(2):165–170, 1994.

[2] M. Balabane and V. Tirel. Décomposition de domaine pour un calcul hybride de l'équation de Helmholtz. *C. R. Acad. Sci. Paris Sér. I Math.*, 324(3):281–286, 1997.

[3] J.-D. Benamou. An introduction to Eulerian geometrical optics (1992-2002). *SIAM J. Sci. Comp.*, 2002. To appear.

[4] J.-D. Benamou, F. Castella, T. Katsaounis, and B. Perthame. High frequency limit of the Helmholtz equations. *Rev. Mat. Iberoamericana*, 18(1):187–209, 2002.

[5] J.-D. Benamou, O. Lafitte, R. Sentis, and I. Solliec. A geometrical optics-based numerical method for high frequency electromagnetic fields computations near fold caustics - part I. *J. Comput. Appl. Math.*, 156:93–125, 2003.

[6] J. Van Bladen. *Electromagnetic fields*. Hemisphere Publishing Corporation, 1985.

[7] D. Bouche, F. Molinet, and R. Mittra. *Asymptotic methods in electromagnetics*. Springer-Verlag, Berlin, 1997. Translated from the 1994 French original by Patricia and Daniel Gogny and revised by the authors.

[8] W. D. Burnside, C. L. Yu, and R. J. Marhefka. A technique to combine the geometric theory of diffraction and the moment method. *IEEE T. Antenn. Propag.*, AP-33:551–558, 1975.

[9] Q. Carayol and F. Collino. Error estimates in the Fast Multipole Method for scattering problems. Part 1: Truncation of the Jacobi–Anger series. *M2AN Math. Model. Numer. Anal.*, 2004.

[10] O. Cessenat and B. Despres. Application of an ultra weak variational formulation of elliptic PDEs to the two-dimensional Helmholtz problem. *SIAM J. Numer. Anal.*, 35(1):255–299 (electronic), 1998.

[11] R. Coifman, V. Rokhlin, and S. Greengard. The fast multipole method for the wave equation: A pedestrian prescription. *IEEE T. Antenn. Propag.*, 35(3):7–12, 1993.

[12] D. Colton and R. Kress. *Inverse acoustic and electromagnetic scattering theory*, volume 93 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, second edition, 1998.

[13] A. de la Bourdonnaye and M. Tolentino. Numerical simulation of scattering problems with Fourier-integral operators. *Math. Models Methods Appl. Sci.*, 7(5):613–631, 1997.

[14] J. J. Duistermaat. Oscillatory integrals, Lagrange immersions and unfolding of singularities. *Comm. Pure Appl. Math.*, 27:207–281, 1974.

[15] W. E and B. Engquist. The heterogeneous multiscale method. *Commun. Math. Sci.*, 1(1):87–132, 2003.

[16] B. Engquist and O. Runborg. Computational high frequency wave propagation. *Acta Numerica*, 12, 2003.

[17] P. Gérard, P. A. Markowich, N. J. Mauser, and F. Poupaud. Homogenization limits and Wigner transforms. *Comm. Pure Appl. Math.*, 50(4):323–379, 1997.

[18] E. Giladi and J. B. Keller. A hybrid numerical asymptotic method for scattering problems. *J. Comput. Phys.*, 174(1):226–247, 2001.

[19] G. H. Golub and C. F. van Loan. *Matrix Computations*. John Hopkins University Press, 1996.

[20] S. Hagdahl. Hybrid methods for computational electromagnetics in the frequency domain. Licentiate's thesis, NADA, KTH, Stockholm, 2003.

[21] U. Jakobus and F. M. Landstorfer. Improved PO-MM hybrid formulation for scattering from three-dimensional perfectly conducting bodies of arbitrary shape. *IEEE T. Antenn. Propag.*, 43(2):162–169, February 1995.

[22] I. G. Kevrekidis, C. W. Gear, J. Hyman, P. G. Kevrekidis, O. Runborg, and C. Theodoropoulos. Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level tasks. *Comm. Math. Sci.*, 1:715–762, 2003.

[23] O. Lafitte. The kernel of the Neumann operator for a strictly diffractive analytic problem. *Commun. Part. Diff. Eq.*, 20(3-4):419–483, 1995.

[24] P.-L. Lions and T. Paul. Sur les mesures de Wigner. *Rev. Mat. Iberoamericana*, 9(3):553–618, 1993.

[25] L. Lorch. Alternative proof of a sharpened form of Bernstein's inequality for Legendre polynomials. *Appl. Anal.*, 14(3):237–240, 1982/83.

[26] L. Lorch. Corrigendum: "Alternative proof of a sharpened form of Bernstein's inequality for Legendre polynomials" [Appl. Anal. **14** (1982/83), no. 3, 237–240; MR 84k:26017]. *Appl. Anal.*, 50(1-2):47, 1993.

[27] J. J. Maciel and L. B. Felsen. Discretized Gabor-based beam algorithm for time-harmonic radiation from two-dimensional truncated planar aperture distributions–I: Formulation and solution. *IEEE T. Antenn. Propag.*, 50:1751–1759, 2002.

[28] J. J. Maciel and L. B. Felsen. Discretized Gabor-based beam algorithm for time-harmonic radiation from two-dimensional truncated planar aperture distributions–II: Asymptotics and numerical tests. *IEEE T. Antenn. Propag.*, 50:1760–1768, 2002.

[29] L. N. Medgyesi-Mitschang and D.-S. Wang. Hybrid methods for analysis of complex scatterers. *P. IEEE*, 77(5):770–779, 1989.

[30] R. Roy and T. Kailath. ESPRIT—estimation of signal parameters via rotationnal invariance techniques. *IEEE T. Acoust. Speech*, 37:984–995, 1989.

[31] R. O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE T. Antenn. Propag.*, 34:276–280, 1986.

[32] J. Song and W. C. Chew. Error analysis for the truncation of multipole expansion of vector Green's function. *IEEE Microw. Wirel. Co.*, 11:311–313, July 2001.

[33] J. Song, C. C. Lu, and W. C. Chew. Numerical accuracy of multipole expansion for 2-D MLFMA. *IEEE Antennas Propag.*, 51(8):1883–1890, August 2003.

[34] G. A. Thiele and T. H. Newhouse. A hybride technique for combining moment methods with a geometrical theory of diffraction. *IEEE T. Antenn. Propag.*, 23:62–69, 1975.