OPTIMIZATION OF THE SHAPE OF GAUSSIAN BEAMS OF A FIXED LENGTH

Luděk Klimeš

Institute of Geology and Geotechnics, Czechosl. Acad. Sci., Prague)*

Резюме: *Предлагается методика определения формы гауссевых пучков, в которой минимизируется данная целевая функция. Точный вид целевой функции не определяется, методику возможно применить для широкого класса целевых функций. Общий вид целевой функции позволяет, например, минимизировать вдоль центрального луча средний квадрат квадратичных изменений комплексного эйконала гауссевого пучка и средний квадрат ширины гауссевого пучка. Предлагаемая методика позволяет даже учесть ошибку, возникающую при переходе гауссевого пучка через структуральную границу раздела. В методике предполагается, что длина гауссевого пучка априорно заданная. Это значит, что длина гауссевого пучка не является свободным параметром в процедуре минимизации. Большинство предложенных в сейсмологической литературе методов для определения формы гауссевых пучков является частным случаем предлагаемой методики.*

Summary: *The procedure of choosing the shape of Gaussian beams in order to minimize a given object function of a certain kind is proposed. The general form of the object function enables both the average square of the quadratic variation of the phase and the average square of the beamwidth to be minimized along the central ray. The error of the transformation of the Gaussian beams at the structural interfaces may also be taken into account. Most of the hitherto published suggestions of how to choose the shape of Gaussian beams are special cases of the described procedure. The aim of this paper is not to propose the object function to be minimized, but only to describe the minimization of a given object function. The minimization assumes the a priori known lengths of the central rays of the Gaussian beams (i.e. the lengths of the beams are not free parameters in the minimization procedure).*

1. INTRODUCTION

Recently, the Gaussian beam summation method has found a large number of applications in seismology and their number is ever increasing.

The great advantage of the Gaussian beam summation method is its flexibility. Gaussian beams may be taken to be narrow or broad, with an arbitrary curvature of the wavefront. Many high-frequency asymptotic methods may be simulated by a specific choice of the shape of beams, and a great variety of results can be obtained with the use of Gaussian beams of various shapes for numerical wavefield modelling. The summation of extremely narrow beams at the receivers yields the ray method, the summation of extremely broad beams at the receivers yields the Chapman-Maslov method [6], [15], etc. The choice of Gaussian beams of finite (nonzero) width makes the computed wavefield continuous. However, such a continuous wavefield may be very similar to the singular wavefield obtained by another asymptotic method (e.g. the ray method) if the shape of the used beams is not sufficiently different from that yielding the mentioned asymptotic method. In such a case, the Gaussian beam method only provides a smoothing of the incorrect results of the wavefield modelling. The accuracy and usefulness of the Gaussian beam summation method thus depends on the proper choice of the shape of the beams.

*) A ddress: V Holešovičkách 41, 182 09 Praha 8, Czechoslovakia.

Unfortunately, at present we have no reliable method of selecting the shape of Gaussian beams. We only guess some of the qualitative criteria for choosing the beam shape, but we are usually not able to express them quantitatively in order to use them in the Gaussian beam programs. We know how to select the shape of beams in some simple structures, but do not have the general algorithm of the choice of the beam shape. Numerical tests and comparisons of the Gaussian beam results with more accurate methods allow us to estimate the optimum shape of the beams only in some special seismic models. Several authors suggested some algorithms of the choice of the shape of Gaussian beams, but none of them is generally applicable.

We feel that the velocity changes in the model should be sufficiently smooth and small within the range of the width of a Gaussian beam. This forces the beams to be narrow. On the other hand, we feel that the complex-valued curvature of a Gaussian beam wavefront should be small in the effective region of the beam, which forces the phasefronts (i.e. the surfaces of the constant real part of the complex-valued phase) of the beams to be plane. This also forces the widths of beams to be large if the beam phasefronts are nearly plane.

At an interface, a Gaussian beam is transformed more precisely if there are smaller differences between the angles of incidence at the interface of the individual complex rays forming the Gaussian beam, within the effective region of the beam. This forces the second derivatives along the interface of the real part of the complex-valued phase (travel time) of the beam to be as small as possible. Simultaneously, the imaginary part of the second differential along the interface of the beam phase is forced to equal the real part. However, it should be sufficiently large for the complex-valued phase of the beam to be approximated by the Taylor expansion up to the second order in the effective region of the beam.

For instance, in a 1-D layered model with homogeneous layers, the plane waves are exact solutions of the elastodynamic equations and may be exactly transformed at structural interfaces. The broadest beams are then expected to yield the best results. The superposition of sufficiently broad Gaussian beams actually allows us to compute, e.g., head waves with great accuracy [11]. However, even in this case, the Gaussian beams should be of a finite width in order to avoid spurious arrivals, well-known from the reflectivity method [5] and caused by an approximate expansion of the source into the outgoing parts of the homogeneous plane waves.

Similarly as in the 1-D layered models with homogeneous layers, in any general 1-D model with a small velocity gradient in the vicinity of the receivers, relatively broad Gaussian beams are recommended [12].

If we consider an interface with an edge to be the limiting case of a bent interface, we are led to choose very narrow Gaussian beams in the vicinity of the edge. This choice may yield good results for the diffracted wave [7]. Unfortunately, the accurate computation of the reflected or transmitted wave needs the beams to be broad as discussed above. The transition from the relatively broad beams in the smooth parts of the interface to the very narrow beams near the edge may cause serious problems with the convergence of the integral superposition of Gaussian beams. It may then be desirable to use beams of a standard width in the superposition even in the vicinity of the edge, and to take into account the corresponding edge wave [9].

Let us consider an alternative situation: The surface or the profile along which Gaussian beams are superposed intersects a structural interface. The amplitude of the superposition of Gaussian beams then decreases near the interface as the beams beyond the interface correspond to the other of the pair of incident-transmitted or reflected-transmitted waves than the considered wave, so that they do not form its analytical continuation. The amplitude decay reaches one half at the interface. Moreover, the superposition of the beams of the wave beyond the interface analytically continues across the interface and forms a spurious wave. We may try to reduce the amplitude decay and the spurious waves by contracting the width of the beams as they approach the interface [16], but the shape of the beams must vary drastically in the vicinity of the interface and the integral superposition of the beams may then diverge. Instead, it may be better to preserve a reasonable width of Gaussian beams even in the vicinity of the interface and, in the superposition, to omit the beams corresponding to the waves beyond the interface. The spurious waves are thus removed. The amplitudes, decreased near the interface, may be restored by incorporating the virtual beams beyond the interface, which correspond to the analytical continuation of the wave under consideration, beyond the interface [13].

Besides the error of the Gaussian beam solution of the elastodynamic equations, also the error of the expansion of the given initial conditions into Gaussian beams must be considered. For instance, in the vicinity of a point source, the Gaussian beams should be narrower than required for the smooth initial conditions (specified, e.g., along a surface) in the same region of the same model. The most detailed study of the expansion error was published in [17]. Unfortunately, the existence and positions of the spurious saddle points in the expansion are strongly dependent on the coordinates in which the phase of the beams is approximated by the Taylor expansion. Since in [17] the beam phase is approximated in the ray-centred coordinate system, whereas the paraxial approximation of the phase in general coordinates is presently most common, we shall not summarize their results here.

In this paper we propose a procedure for determining the shape of Gaussian beams so that they minimize the integral of a certain expression along a fixed part of the beam's central ray. The general form of the minimized expression is introduced in the next section. This approach is close to that of [8], but the algorithm proposed here is more general. The expressions under consideration allow us to minimize not only the width of the beams, but also the quadratic variations of the complex-valued phase along an arbitrary surface, e.g., along a structural interface or along a wavefront tangent plane.

The reader should be familiar with the Gaussian beam method and with the properties of the "ray propagator matrix". An ample list of references related to Gaussian beams is given in [2]. The basic properties of the ray propagator matrix are summarized in [3].

2. FORMULATION

2.1 The specification of some used quantities

The matrices will be denoted concurrently by boldface letters (e.g. $\mathbf{A}$) and by means of their components (e.g. A_{KL}). In the case of the component notation, the capital-letter indices take the values $K, L, \ldots = 1, 2$; the lower-case indices take the values $k, l, \ldots = 1, 2, 3$; the Greek indices take the values $\alpha, \beta, \ldots = 1, 2, 3, 4$. The dagger (e.g. $\mathbf{A}^+$) denotes the Hermitian adjoint (transpose for real-valued matrices).

We denote by A_k the amplitude and by ϑ the complex-valued phase of a frequency-domain Gaussian beam

$$g_k = A_k \exp(\mathrm{i}\omega\vartheta) . \tag{1}$$

Here ω is the positive circular frequency. In ray-centred coordinates q_k, where q_3 is an independent variable along the ray, the quadratic Taylor expansion of the phase has the form of

$$\vartheta(q_j) = \tau(q_3) + \tfrac{1}{2} q_K M_{KL}(q_3)\, q_L , \tag{2}$$

where τ is the travel time along the central ray. The phase may also be expressed in any other coordinate system. In such a case we always restrict the Taylor expansion up to the quadratic terms.

The second differential $\mathbf{M}$ of the phase along the plane q_1q_2 tangent to the phasefront is a 2×2 matrix with a symmetric real part $\mathbf{R}$ and a positive-definite symmetric imaginary part $\mathbf{Y}$,

(3) $$\mathbf{M} = \mathbf{R} + \mathrm{i}\mathbf{Y} .$$

It may be expressed as

(4) $$\mathbf{M} = \mathbf{P}\mathbf{Q}^{-1} ,$$

where we have denoted by

(5) $$\begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} = \begin{pmatrix} \mathbf{1} \\ \mathbf{M} \end{pmatrix} \mathbf{Q}$$

the solution of the dynamic ray-tracing system

(6) $$\frac{\mathrm{d}}{v^2\,\mathrm{d}\tau} \begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} = \Sigma \begin{pmatrix} \mathbf{V}v^{-3} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix},$$

corresponding to the Gaussian beam. Here

(7) $$\Sigma = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix}$$

is the 4×4 antisymmetric matrix, $\mathbf{0}$ and $\mathbf{1}$ being zero and identity 2×2 matrices. The 2×2 matrix $\mathbf{V}$ in (6) is the second differential of the propagation velocity along the phasefront tangent plane.

Any solution of the dynamic ray-tracing system (6) with the initial conditions

(8) $$\begin{pmatrix} \mathbf{Q}(q_3^{(0)}) \\ \mathbf{P}(q_3^{(0)}) \end{pmatrix} = \begin{pmatrix} \mathbf{Q}_0 \\ \mathbf{P}_0 \end{pmatrix} = \begin{pmatrix} \mathbf{1} \\ \mathbf{M}_0 \end{pmatrix} \mathbf{Q}_0$$

may be written in the form

(9) $$\begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} = \Pi \begin{pmatrix} \mathbf{Q}_0 \\ \mathbf{P}_0 \end{pmatrix},$$

where the ray propagator matrix

(10) $$\Pi(q_3, q_3^{(0)}) = \begin{pmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \\ \mathbf{P}_1 & \mathbf{P}_2 \end{pmatrix}$$

is the fundamental 4×4 matrix of the solutions of the dynamic ray-tracing system (6) with the unit initial conditions

(11) $$\Pi(q_3^{(0)}, q_3^{(0)}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

at point $q_3 = q_3^{(0)}$. Since matrix

(12) $$\begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix}^{+} \Sigma \begin{pmatrix} \mathbf{Q}' \\ \mathbf{P}' \end{pmatrix}$$

is constant along the ray for any two solutions

$$\begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \mathbf{Q}' \\ \mathbf{P}' \end{pmatrix} \tag{13}$$

of the dynamic ray tracing system (6), the ray propagator matrix $\mathbf{\Pi}$ is symplectic:

$$\mathbf{\Pi}^{+}\mathbf{\Sigma}\mathbf{\Pi} = \mathbf{\Sigma}\,. \tag{14}$$

2.2 Notes and speculations

Let us consider a Gaussian beam incident at a structural interface specified by the equation $f(x_k) = 0$. At the point of incidence, let us introduce the local Cartesian coordinates z_k, defined by means of the orthonormal basis vectors Z_{j1}, Z_{j2}, Z_{j3},

$$Z_{jk} = \partial x_j / \partial z_k\,, \tag{15}$$

where

$$Z_{j3} = \partial f / \partial x_j (\partial f / \partial x_k\, \partial f / \partial x_k)^{-1/2} \tag{16}$$

is the unit normal to the interface. Here x_k are the general Cartesian coordinates. The interface in the vicinity of the point of incidence may be parametrized by the coordinates z_1, z_2. The complex-valued phase of the Gaussian beam along the interface is ([10], equation 45)

$$\vartheta(z_1, z_2) = \tau + p_j Z_{jK} z_K + \tfrac{1}{2} z_K z_L (H_{KM} M_{MN} H_{LN} + E_{KL})\,, \tag{17}$$

where

$$H_{jk} = \partial z_j / \partial q_k = Z_{nj}\, \partial x_n / \partial q_k \tag{18}$$

is the transformation matrix from ray-centred coordinates q_j to coordinates z_j. Matrix E_{AB} depends on the slowness and its gradient, on the curvature of the interface and on the angle of incidence. The linear variations (with respect to z_K) of the slowness vector components p_K tangent to the interface are

$$\delta p_K = (H_{KM} M_{MN} H_{LN} + E_{KL})\, z_L\,, \tag{19}$$

and the linear variation of any reflection/transmission coefficient $R = R(p, \alpha, \beta, \varrho)$ is

$$\delta R = \frac{\partial R}{\partial p} \frac{p_K}{p} \delta p_K + \left(\frac{\partial R}{\partial \alpha} \frac{\partial \alpha}{\partial z_L} + \frac{\partial R}{\partial \beta} \frac{\partial \beta}{\partial z_L} + \frac{\partial R}{\partial \varrho} \frac{\partial \varrho}{\partial z_L} \right) z_L\,, \tag{20}$$

where $p = (p_K p_K)^{1/2}$ is the ray parameter. This equation may be expressed, similarly as (19), in the form of

$$\delta R = W_K (M_{KM} - K_{KM})\, H_{LM} z_L\,, \tag{21}$$

where W_K, K_{KM} are the functions of the angle of incidence, of the medium parameters and the interface curvature. The average square

$$\langle |\delta R|^2 \rangle = \iint |\delta R \exp(\mathrm{i}\omega\vartheta)|^2\, \mathrm{d}z_1\, \mathrm{d}z_2 / \{\iint |\exp(\mathrm{i}\omega\vartheta)|^2\, \mathrm{d}z_1\, \mathrm{d}z_2\} \tag{22}$$

of linear variation (21) of the reflection/transmission coefficient is

$$\langle|\delta R|^2\rangle = \tfrac{1}{2}\omega^{-1}\,\mathrm{tr}[\mathbf{W}(\mathbf{M} - \mathbf{K})\,\mathbf{Y}^{-1}(\mathbf{M} - \mathbf{K})^{+}]\,, \tag{23}$$

where

$$W_{KL} = W_K W_L\,. \tag{24}$$

Any sum of the expressions of the kind of (23) may be expressed in the form of (23), with $\mathbf{W}$ being the sum of the particular weighting functions (24). We feel that some expressions of the kind of (23) should be minimized in order to minimize the error caused by the transformation of the Gaussian beams across the interface. Note that also the relative average square

$$\langle v^2|\delta p_K\,\delta p_K|\rangle = \iint v^2|\delta p_K\,\delta p_K|\,|\exp(i\omega\vartheta)|^2\,dz_1\,dz_2\,. \left\{\iint|\exp(i\omega\vartheta)|^2\,dz_1\,dz_2\right\}^{-1} \tag{25}$$

of the linear variations (19) of the slowness vector components tangent to an arbitrary curved surface may be written in the form of (23),

$$\langle v^2|\delta p_K\,\delta p_K|\rangle = \tfrac{1}{2}\omega^{-1}v^2\,\mathrm{tr}\,[\mathbf{H}^{+}\mathbf{H}(\mathbf{M} - \mathbf{K})\,\mathbf{Y}^{-1}(\mathbf{M} - \mathbf{K})^{+}]\,, \tag{26}$$

and may be minimized. Roughly speaking, expression (26) represents a measure of the quadratic variations of the beam phase along a curved surface within the effective region of the beam.

Kim and Garmany ([8], condition 1) suggested to minimize the quadratic variation of the phase along the wavefront tangent plane. The complex-valued curvature of the phasefront of the Gaussian beam may be minimized within the effective region of the beam in the following sense: The quadratic variation of the complex-valued phase may be expressed in terms of the linear variation

$$\delta p_K^{(q)} = M_{KL}q_L \tag{27}$$

of the ray-centred components of the complex-valued slowness vector along the phase-front tangent plane. The relative average square of the components $\delta p_K^{(q)}$ of the slowness vector, perpendicular to the central ray, is

$$\langle v^2|\delta p_K^{(q)}\,\delta p_K^{(q)}|\rangle = \tfrac{1}{2}\omega^{-1}v^2\,\mathrm{tr}\,[\mathbf{M}\mathbf{Y}^{-1}\mathbf{M}^{+}]\,, \tag{28}$$

which is a special case of (26). Expression (28) also represents the average square of the linear variation of the normal vp_j, to the phasefront. Let us mention that the linear variation of the normal vp_j to the phasefront is closely related to the additional components of the beam amplitude.

The effect of a smooth heterogeneity of a Gaussian velocity profile on the Gaussian beam synthetic seismograms was studied in [14]. For a velocity contrast of 10% lower than the background, they suggested that the halfwidths L of Gaussian beams should be approximately the same or smaller than the radius (halfwidth) l of the heterogeneity,

$$L \leqq l\,. \tag{29}$$

In three dimensions, the halfwidth of a Gaussian beam is a vectorial distance L_K from the central ray to a point on the boundary of the beam spot ellipse,

$$\tfrac{1}{2}\omega L_K Y_{KL} L_L = 1\,. \tag{30}$$

Let us express condition (29) in terms of a velocity gradient. We introduce the 2×2 positive-semidefinite matrix U_{KL} as the supremum of the matrices

$$v^{-2}(\partial v/\partial q_K)\,(\partial v/\partial q_L) \tag{31}$$

over the spot ellipse, see (30). Note that in the terminology of [1], the norm of the vector $\frac{1}{2}v^{-1}\,\partial v/\partial q_k$ corresponds to the inverse characteristic length of the medium, and matrix $\frac{1}{4}\mathbf{U}$ may thus be considered to correspond to the inverse square of the characteristic length of the medium. For the 10% Gaussian inhomogeneity we have

$$\mathbf{U} = \mathbf{1}/(120 l^2) \tag{32}$$

and inequality (29) may be expressed as

$$\omega^{-1}\mathbf{U}^{1/2}\mathbf{Y}^{-1}\mathbf{U}^{1/2} \leqq \mathbf{1}/240\,, \tag{33}$$

see (30). This condition restricts the variation of the slowness over the spot ellipse. Alternatively, it restricts the relative variation of the norm $(p_k p_k)^{1/2} \approx p_3^{(q)}$ of the slowness vector over the spot ellipse. Condition (33) forces the Gaussian beams to be narrow, which conforms to [8], condition 2.

Seven validity conditions for Gaussian beams were proposed in [1]. Four of them depend on the shape of Gaussian beams: *The far-field condition*

$$\omega^{-1}|A|^{-1}\,|\mathrm{d}A/\mathrm{d}\tau| \leqq 1/13 \tag{34}$$

restricts the relative variation of the scalar beam amplitude A per wavelength along the central ray. *The ray-centred coordinate system regularity condition* requires the ray-receiver distance q_K (expressed in the ray-centred coordinate system) to be sufficiently small with respect to the curvature $-v^{-1}\,\partial v/\partial q_K$ fo the ray,

$$|q_K v^{-1}\,\partial v/\partial q_K| < 1/5 \tag{35}$$

for the central ray and

$$q_K U_{KL} q_L < 1/25 \tag{36}$$

if we consider all rays within the spot ellipse. *The paraxial condition*

$$v|q_K M_{KM} M_{ML} q_L|^{1/2} < 1/4 \tag{37}$$

requires the ray-receiver distance q_K to be sufficiently small with respect to the curvature $-v\mathbf{M}$ of the phasefront. Condition (37) was proposed in [1] in a slightly different form, with the real part $-v\mathbf{R}$ of the phasefront curvature instead of $-v\mathbf{M}$. However, we believe that the whole complex-valued phasefront curvature $-v\mathbf{M}$ should be used in (37). The last proposed condition is *the Fresnel 2-condition*

$$\tfrac{1}{4}\omega^{-1}\mathbf{U}^{1/2}|\mathbf{M}|^{-1}\,\mathbf{U}^{1/2} \leqq \mathbf{1}/3\,. \tag{38}$$

All the limits on the r.h.s. of the validity conditions are only informative and were obtained empirically in [1] for the maximum relative error of 20% in the synthetic seismograms. Substituting the beam halfwidths L_K, see (30), for the maximum ray-receiver distance q_K in (36) and (37), we arrive at

$$\omega^{-1}\mathbf{U}^{1/2}\mathbf{Y}^{-1}\mathbf{U}^{1/2} < \mathbf{1}/50\,, \tag{39}$$

which is a condition analogous to (33), and at

$$\omega^{-1}v^2\mathbf{M}\mathbf{Y}^{-1}\mathbf{M}^+ < \mathbf{1}/32\,. \tag{40}$$

The necessary and sufficient conditions for the inequality (39) are

$$\omega^{-1}\operatorname{tr}(\mathbf{U}\mathbf{Y}^{-1}) < 1/25 \quad \text{and} \quad \omega^{-1}\operatorname{tr}(\mathbf{U}\mathbf{Y}^{-1}) < 1/50\,, \tag{41}$$

the necessary and sufficient conditions for the inequality (40) are

$$\omega^{-1}v^2\operatorname{tr}(\mathbf{M}\mathbf{Y}^{-1}\mathbf{M}^+) < 1/16 \quad \text{and} \quad \omega^{-1}v^2\operatorname{tr}(\mathbf{M}\mathbf{Y}^{-1}\mathbf{M}^+) < 1/32\,. \tag{42}$$

These inequalities may be considered to be quantified versions of the condition to minimize (28). Since

$$\mathrm{d}A/\mathrm{d}\tau = -\tfrac{1}{2}Av^2\operatorname{tr}\mathbf{M}\,, \tag{43}$$

the far-field condition (34) may be rewritten to read

$$\omega^{-1}v^2|\operatorname{tr}\mathbf{M}| \leqq 2/13\,, \tag{44}$$

and it is satisfied in consequence of satisfying condition (42) because

$$|\operatorname{tr}\mathbf{M}| \leqq \operatorname{tr}(\mathbf{M}\mathbf{Y}^{-1}\mathbf{M}^+)\,. \tag{45}$$

Likewise, the Fresnel 2-condition (38) is satisfied in consequence of satisfying condition (39) because

$$|\mathbf{M}| \geqq \mathbf{Y}\,. \tag{46}$$

Any of the expressions appearing in (23), (26), (28), (41), (42) have the general form of

$$\omega^{-1}\operatorname{tr}\{\mathbf{U}\mathbf{Y}^{-1} + \mathbf{W}(\mathbf{M}-\mathbf{K})\,\mathbf{Y}^{-1}(\mathbf{M}-\mathbf{K})^+\}\,, \tag{47}$$

where the 2×2 matrices $\mathbf{U}(q_3)$, $\mathbf{W}(q_3)$, $\mathbf{K}(q_3)$ depend on the parameter q_3 along the central ray of the Gaussian beam.

2.3 Object function

Assume that the dependence of the error of the Gaussian beam solution of the elastodynamic equations on the shape of the Gaussian beam may be roughly expressed in terms of the object function of the form, see (47),

$$\begin{aligned} T &= \int_{q_3^{(1)}}^{q_3^{(2)}} \operatorname{tr}\{\mathbf{U}\mathbf{Y}^{-1} + \mathbf{W}(\mathbf{M}-\mathbf{K})\,\mathbf{Y}^{-1}(\mathbf{M}-\mathbf{K})^+\}\,\mathrm{d}q_3 \\ &= \int_{q_3^{(1)}}^{q_3^{(2)}} \operatorname{tr}\{\mathbf{U}\mathbf{Y}^{-1} + \mathbf{W}[(\mathbf{R}-\mathbf{K})\,\mathbf{Y}^{-1}(\mathbf{R}-\mathbf{K})^+ + \mathbf{Y}]\}\,\mathrm{d}q_3\,, \end{aligned} \tag{48}$$

where $\mathbf{U}(q_3)$ and $\mathbf{W}(q_3)$ are real symmetric positive-semidefinite 2×2 matrices and $\mathbf{K}(q_3)$ is a real 2×2 matrix, for $\mathbf{M}(q_3)$, $\mathbf{R}(q_3)$ and $\mathbf{Y}(q_3)$ see (3). Note that the weighting functions $\mathbf{U}(q_3)$, $\mathbf{W}(q_3)$ may be of a delta type at some points, e.g., at interfaces. It is convenient to represent the weighting functions $\mathbf{U}(q_3)$, $\mathbf{W}(q_3)$, $\mathbf{K}(q_3)$ in terms of the 4×4 real symmetric positive-semidefinite weighting matrix $\mathbf{F}(q_3)$ composed of the 2×2 submatrices $\mathbf{F}_{KL}(q_3)$ as follows

(49) $$\mathbf{F} = \begin{pmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{U} + \mathbf{K}^+\mathbf{W}\mathbf{K} & -\mathbf{K}^+\mathbf{W} \\ -\mathbf{W}\mathbf{K} & \mathbf{W} \end{pmatrix}.$$

For

(50) $$\mathbf{U} = \mathbf{F}_{11} - \mathbf{F}_{12}\mathbf{F}_{22}^{-1}\mathbf{F}_{21}\,, \quad \mathbf{K} = -\mathbf{F}_{22}^{-1}\mathbf{F}_{21}\,, \quad \mathbf{W} = \mathbf{F}_{22}\,,$$

the object function (48) may be expressed in the form

(51) $$T(\mathbf{F}) = \int_{q_3^{(1)}}^{q_3^{(2)}} \operatorname{tr}\{\mathbf{F}_{11}\mathbf{Y}^{-1} + \mathbf{F}_{12}\mathbf{M}\mathbf{Y}^{-1} + \mathbf{F}_{21}\mathbf{Y}^{-1}\mathbf{M}^+ + \mathbf{F}_{22}\mathbf{M}\mathbf{Y}^{-1}\mathbf{M}^+\}\,\mathrm{d}q_3\,,$$

or in the more compact form

(52) $$T(\mathbf{F}) = \int_{q_3^{(1)}}^{q_3^{(2)}} \operatorname{tr}(\mathbf{F}\mathbf{\Psi})\,\mathrm{d}q_3\,,$$

where the 4×4 matrix

(53) $$\mathbf{\Psi} = \begin{pmatrix} \mathbf{1} \\ \mathbf{M} \end{pmatrix} \mathbf{Y}^{-1}(\mathbf{1}, \mathbf{M}^+)$$

is formed of its 2×2 submatrices in the following way:

(54) $$\mathbf{\Psi} = \begin{pmatrix} \mathbf{Y}^{-1} & \mathbf{Y}^{-1}\mathbf{M}^+ \\ \mathbf{M}\mathbf{Y}^{-1} & \mathbf{M}\mathbf{Y}^{-1}\mathbf{M}^+ \end{pmatrix} = \begin{pmatrix} \mathbf{Y}^{-1} & \mathbf{Y}^{-1}\mathbf{R} \\ \mathbf{R}\mathbf{Y}^{-1} & \mathbf{R}\mathbf{Y}^{-1}\mathbf{R} + \mathbf{Y} \end{pmatrix} - \mathrm{i}\mathbf{\Sigma}\,.$$

Since $\mathbf{F}$ is a real symmetric matrix and $\mathbf{\Psi}$ is a self-adjoint matrix, equation (52) reads

(55) $$T(\mathbf{F}) = \int_{q_3^{(1)}}^{q_3^{(2)}} \operatorname{tr}\{\mathbf{F}\,\mathrm{Re}(\mathbf{\Psi})\}\,\mathrm{d}q_3\,,$$

where

(56) $$\mathrm{Re}(\mathbf{\Psi}) = \begin{pmatrix} \mathbf{Y}^{-1} & \mathbf{Y}^{-1}\mathbf{R} \\ \mathbf{R}\mathbf{Y}^{-1} & \mathbf{R}\mathbf{Y}^{-1}\mathbf{R} + \mathbf{Y} \end{pmatrix},$$

see (54). The real symmetric positive-definite 4×4 matrix $\mathrm{Re}(\mathbf{\Psi})$ is symplectic,

(57) $$\mathrm{Re}(\mathbf{\Psi})^+\mathbf{\Sigma}\,\mathrm{Re}(\mathbf{\Psi}) = \mathbf{\Sigma}\,.$$

On the other hand, any real symmetric positive-definite symplectic 4×4 matrix $\mathrm{Re}(\mathbf{\Psi})$ may be written in the form (56) where $\mathbf{R}$ and $\mathbf{Y}$ are real symmetric 2×2 matrices, $\mathbf{Y}$ being a positive-definite matrix. The functional (52), (55) is linear with respect to the weighting functions $\mathbf{F}$,

(58) $$T(a^{(1)}\mathbf{F}^{(1)} + a^{(2)}\mathbf{F}^{(2)}) = a^{(1)}T(\mathbf{F}^{(1)}) + a^{(2)}T(\mathbf{F}^{(2)})\,.$$

Assume that the weighting matrix $\mathbf{F}(q_3)$ is independent of the shape of the Gaussian beam. This assumption is violated for condition (41) with $\mathbf{U}$ being the supremum of (31) over the spot ellipse, but we hope that the assumption can usually be satisfied by replacing the spot ellipse by another reasonable region for evaluating the supremum.

We assume that the expansion error of the given wavefield into Gaussian beams is also involved in (55) and we shall minimize (55) for a single Gaussian beam between the points $q_3 = q_3^{(1)}$ and $q_3 = q_3^{(2)}$ by adjusting the three independent complex-valued components of the matrix $\mathbf{M}(q_3^{(0)})$ at the given point $q_3^{(0)}$. We think that this assumption of the fixed beam length is the most severe restriction of the proposed algorithm since the minimum value of the functional (55) may be too large for a single beam between points $q_3^{(1)}$ and $q_3^{(2)}$. In such a case, a smaller value of the functional (55) may be obtained by segmenting the interval $\langle q_3^{(1)}, q_3^{(2)}\rangle$ into several pieces and minimizing (55) in parts for individual shorter beams even if the error of the superposition of the old beams and the consecutive approximate expansion into the new beams is added at each break point of the interval $\langle q_3^{(1)}, q_3^{(2)}\rangle$.

3. MINIMIZATION OF THE OBJECT FUNCTION

3.1 Expression of the object function in terms of the initial value $\mathbf{M}_0$ of matrix $\mathbf{M}$

We shall rewrite (55), which is the same as (48), (51) or (52), in terms of the initial value

(59) $$\mathbf{M}_0 \equiv \mathbf{M}(q_3^{(0)})$$

of matrix $\mathbf{M}(q_3)$ at an arbitrarily chosen point $q_3^{(0)}$ of the central ray. We denote its real part $\mathbf{R}_0$ and its imaginary part $\mathbf{Y}_0$,

(60) $$\mathbf{M}_0 = \mathbf{R}_0 + i\mathbf{Y}_0 .$$

In view of (12), the matrix

(61) $$\mathbf{Q}^+\mathbf{Y}\mathbf{Q} = (2i)^{-1}\mathbf{Q}^+(\mathbf{M} - \mathbf{M}^+)\,\mathbf{Q} = (2i)^{-1}(\mathbf{Q}^+\mathbf{P} - \mathbf{P}^+\mathbf{Q}) = \\ = (2i)^{-1}\begin{pmatrix}\mathbf{Q}\\ \mathbf{P}\end{pmatrix}^+ \Sigma \begin{pmatrix}\mathbf{Q}\\ \mathbf{P}\end{pmatrix}$$

is constant along the ray,

(62) $$\mathbf{Q}^+\mathbf{Y}\mathbf{Q} = \mathbf{Q}_0^+\mathbf{Y}_0\mathbf{Q}_0 .$$

Here we have put

(63) $$\mathbf{Q}_0 \equiv \mathbf{Q}(q_3^{(0)})\,, \quad \mathbf{P}_0 \equiv \mathbf{P}(q_3^{(0)})\,.$$

Inserting (9) and (62) into (53) written in the form

(64) $$\Psi = \begin{pmatrix}\mathbf{Q}\\ \mathbf{P}\end{pmatrix}(\mathbf{Q}^+\mathbf{Y}\mathbf{Q})^{-1}\begin{pmatrix}\mathbf{Q}\\ \mathbf{P}\end{pmatrix}^+ ,$$

we obtain

(65) $$\Psi = \Pi\begin{pmatrix}\mathbf{Q}_0\\ \mathbf{P}_0\end{pmatrix}(\mathbf{Q}_0^+\mathbf{Y}_0\mathbf{Q}_0)^{-1}\begin{pmatrix}\mathbf{Q}_0\\ \mathbf{P}_0\end{pmatrix}^+ \Pi^+ ,$$

which reads

(66) $$\mathbf{\Psi} = \mathbf{\Pi}\mathbf{\Psi}_0\mathbf{\Pi}^+ ,$$

where

(67) $$\mathbf{\Psi}_0 = \begin{pmatrix} \mathbf{1} \\ \mathbf{M}_0 \end{pmatrix} \mathbf{Y}_0^{-1}(\mathbf{1}, \mathbf{M}_0^+) .$$

The real part of (66) is

(68) $$\mathrm{Re}(\mathbf{\Psi}) = \mathbf{\Pi}\ \mathrm{Re}(\mathbf{\Psi}_0)\ \mathbf{\Pi}^+ ,$$

where

(69) $$\mathrm{Re}(\mathbf{\Psi}_0) = \begin{pmatrix} \mathbf{Y}_0^{-1} & \mathbf{Y}_0^{-1}\mathbf{R}_0 \\ \mathbf{R}_0\mathbf{Y}_0^{-1} & \mathbf{R}_0\mathbf{Y}_0^{-1}\mathbf{R}_0 + \mathbf{Y}_0 \end{pmatrix} .$$

Considering (68), the object function (55), which is the same as (48), (51) and (52), may be written as

(70) $$T(\mathbf{F}) = \mathrm{tr}\{\mathbf{B}(\mathbf{F})\ \mathrm{Re}(\mathbf{\Psi}_0)\} ,$$

where

(71) $$\mathbf{B}(\mathbf{F}) = \int_{q_3(1)}^{q_3(2)} \mathbf{\Pi}^+(q_3, q_3^{(0)})\ \mathbf{F}(q_3)\ \mathbf{\Pi}(q_3, q_3^{(0)})\ \mathrm{d}q_3 .$$

Note that integrals of the kind of (71) also play an important role in the ray perturbation theory ([4], equation 16).

3.2 Minimization of the object function

We shall determine matrix (69), composed of the real symmetric 2×2 matrix $\mathbf{R}_0$ and of the real positive-definite symmetric 2×2 matrix $\mathbf{Y}_0$, minimizing (70) which is equivalent to (55). In other words, we shall find the real positive-definite symmetric symplectic 4×4 matrix $\mathrm{Re}(\mathbf{\Psi}_0)$ minimizing (70).

By decomposing the real positive-definite symmetric 4×4 matrix (71) into 2×2 submatrices

(72) $$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix} ,$$

the object function (70) may be expressed as

(73) $$T = \mathrm{tr}\{\mathbf{B}_{11}\mathbf{Y}_0^{-1} + \mathbf{B}_{12}\mathbf{R}_0\mathbf{Y}_0^{-1} + \mathbf{B}_{21}\mathbf{Y}_0^{-1}\mathbf{R}_0 + \mathbf{B}_{22}(\mathbf{R}_0\mathbf{Y}_0^{-1}\mathbf{R}_0 + \mathbf{Y}_0)\} .$$

It is obvious that this object function has just one local extreme which is simultaneously the global minimum. Note that $\mathbf{B}_{11}$ and $\mathbf{B}_{22}$ are real positive-definite symmetric 2×2 matrices and

(74) $$\mathbf{B}_{21} = \mathbf{B}_{12}^+ .$$

Differentiating (73) with respect to the real symmetric matrix $\mathbf{R}_0$ and putting the result to equal zero, we obtain

(75) $$\mathbf{Y}_0^{-1}\mathbf{B}_{12} + \mathbf{B}_{21}\mathbf{Y}_0^{-1} + \mathbf{Y}_0^{-1}\mathbf{R}_0\mathbf{B}_{22} + \mathbf{B}_{22}\mathbf{R}_0\mathbf{Y}_0^{-1} = \mathbf{0} ,$$

and differentiating (73) with respect to the real symmetric matrix $\mathbf{Y}_0$ and putting the result equal to zero, we obtain

(76) $$-\mathbf{Y}_0^{-1}[\mathbf{B}_{11} + \mathbf{B}_{12}\mathbf{R}_0 + \mathbf{R}_0\mathbf{B}_{21} + \mathbf{R}_0\mathbf{B}_{22}\mathbf{R}_0]\,\mathbf{Y}_0^{-1} + \mathbf{B}_{22} = \mathbf{0}\,.$$

First we solve (75) for fixed $\mathbf{Y}_0$. Equation (75) reads

(77) $$\mathbf{Y}_0^{-1}(\mathbf{B}_{12} + \mathbf{R}_0\mathbf{B}_{22}) + [\mathbf{Y}_0^{-1}(\mathbf{B}_{12} + \mathbf{R}_0\mathbf{B}_{22})]^+ = \mathbf{0}\,,$$

which means that matrix $\mathbf{Y}_0^{-1}(\mathbf{B}_{12} + \mathbf{R}_0\mathbf{B}_{22})$ must be antisymmetric

(78) $$\mathbf{Y}_0^{-1}(\mathbf{B}_{12} + \mathbf{R}_0\mathbf{B}_{22}) = r\boldsymbol{\sigma}\,,$$

where

(79) $$\boldsymbol{\sigma} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

The following relations are valid for matrix $\boldsymbol{\sigma}$ and any 2×2 real matrix $\mathbf{A}$:

(80) $$\boldsymbol{\sigma}\mathbf{A}^+\boldsymbol{\sigma} = -\mathbf{A}^{-1}\det\mathbf{A}\,,$$

(81) $$\boldsymbol{\sigma}\mathbf{A} + \mathbf{A}^+\boldsymbol{\sigma} = \boldsymbol{\sigma}\,\mathrm{tr}\,\mathbf{A}\,,$$

(82) $$\mathbf{A}^+\boldsymbol{\sigma}\mathbf{A} = \boldsymbol{\sigma}\det\mathbf{A}\,.$$

Equation (78) yields

(83) $$\mathbf{R}_0 = -\mathbf{B}_{12}\mathbf{B}_{22}^{-1} + r\mathbf{Y}_0\boldsymbol{\sigma}\mathbf{B}_{22}^{-1}\,.$$

Matrices $\mathbf{R}_0$ and, consequently, $\mathbf{B}_{22}\mathbf{R}_0\mathbf{B}_{22}$ must be symmetric

(84) $$(\mathbf{B}_{22}\mathbf{R}_0\mathbf{B}_{22}) - (\mathbf{B}_{22}\mathbf{R}_0\mathbf{B}_{22})^+ = \mathbf{0}\,,$$

and, therefore,

(85) $$-\mathbf{B}_{22}\mathbf{B}_{12} + r\mathbf{B}_{22}\mathbf{Y}_0\boldsymbol{\sigma} + \mathbf{B}_{21}\mathbf{B}_{22} + r\boldsymbol{\sigma}\mathbf{Y}_0\mathbf{B}_{22} = \mathbf{0}\,.$$

Applying (81) to (85) we obtain

(86) $$r\boldsymbol{\sigma} = (\mathbf{B}_{22}\mathbf{B}_{12} - \mathbf{B}_{21}\mathbf{B}_{22})\,[\mathrm{tr}(\mathbf{Y}_0\mathbf{B}_{22})]^{-1}\,,$$

which may be inserted into (83) to arrive at

(87) $$\mathbf{R}_0 = \mathbf{X} - (\mathbf{Y}_0\mathbf{B}_{22})(\mathbf{X} - \mathbf{X}^+)\,[\mathrm{tr}(\mathbf{Y}_0\mathbf{B}_{22})]^{-1}\,,$$

where

(88) $$\mathbf{X} = -\mathbf{B}_{12}\mathbf{B}_{22}^{-1}\,.$$

Equation (76) may be written in the form

(89) $$-\mathbf{Y}_0^{-1}[\mathbf{C}_{11} + (\mathbf{R}_0 + \mathbf{B}_{12}\mathbf{B}_{22}^{-1})\,\mathbf{B}_{22}(\mathbf{R}_0 + \mathbf{B}_{22}^{-1}\mathbf{B}_{21})]\,\mathbf{Y}_0^{-1} + \mathbf{B}_{22} = \mathbf{0}\,,$$

where

(90) $$\mathbf{C}_{11} = \mathbf{B}_{11} - \mathbf{B}_{12}\mathbf{B}_{22}^{-1}\mathbf{B}_{21}\,.$$

Inserting (87) into (89) we obtain

(91) $$-\mathbf{Y}_0^{-1}\mathbf{C}_{11}\mathbf{Y}_0^{-1} + \mathbf{B}_{22}(\mathbf{X} - \mathbf{X}^+)\,\mathbf{B}_{22}(\mathbf{X} - \mathbf{X}^+)\,\mathbf{B}_{22}[\mathrm{tr}(\mathbf{Y}_0\mathbf{B}_{22})]^{-2} + \\ + \mathbf{B}_{22} = \mathbf{0}\,,$$

which may, using (80) or (82), be turned into

(92) $$-\mathbf{Y}_0^{-1}\mathbf{C}_{11}\mathbf{Y}_0^{-1} + y\mathbf{B}_{22} = \mathbf{0},$$
where
(93) $$y = 1 - \det(\mathbf{X} - \mathbf{X}^+)\det(\mathbf{B}_{22})[\mathrm{tr}(\mathbf{Y}_0\mathbf{B}_{22})]^{-2}.$$
Equation (92) reads
(94) $$(\mathbf{B}_{22}^{1/2}\mathbf{Y}_0\mathbf{B}_{22}^{1/2})^2 = y^{-1}\mathbf{B}_{22}^{1/2}\mathbf{C}_{11}\mathbf{B}_{22}^{1/2}$$
and, consequently,
(95) $$\mathbf{B}_{22}^{1/2}\mathbf{Y}_0\mathbf{B}_{22}^{1/2} = y^{-1/2}\mathbf{S},$$
where
(96) $$\mathbf{S} = (\mathbf{B}_{22}^{1/2}\mathbf{C}_{11}\mathbf{B}_{22}^{1/2})^{1/2}.$$

Note that the square root of the positive-semidefinite self-adjoint matrix $\mathbf{B}$ is the positive-semidefinite self-adjoint matrix $\mathbf{A} = \mathbf{B}^{1/2}$ satisfying

(97) $$\mathbf{AA} = \mathbf{B}.$$

Equation (95) may be inserted into (93) to arrive at

(98) $$y = 1 - \det(\mathbf{X} - \mathbf{X}^+)\det(\mathbf{B}_{22})[\mathrm{tr}(\mathbf{S})]^{-2}\,y,$$

which together with (95) yields

(99) $$\mathbf{Y}_0 = \mathbf{B}_{22}^{-1/2}\mathbf{S}\mathbf{B}_{22}^{-1/2}\{1 + \det(\mathbf{X} - \mathbf{X}^+)\det(\mathbf{B}_{22})[\mathrm{tr}(\mathbf{S})]^{-2}\}^{1/2}.$$

The object function (48), (51), (52) has its minimum value for the choice $\mathbf{M}(q_3^{(0)}) = \mathbf{R}_0 + \mathrm{i}\mathbf{Y}_0$. Matrix $\mathbf{Y}_0$ may be evaluated using (88), (90), (96) and (99). Matrix $\mathbf{R}_0$ is given by (87), where $\mathbf{Y}_0$ is the result of (99). Matrix $\mathbf{B}$, see (72), is defined by (71).

3.3 Minimum value of the object function

For $\mathbf{R}_0$, $\mathbf{Y}_0$ satisfying (76), equation (73) is reduced to

(100) $$T_{\mathrm{MIN}} = 2\,\mathrm{tr}(\mathbf{Y}_0\mathbf{B}_{22}).$$

This yields, after the insertion of (99),

(101) $$T_{\mathrm{MIN}} = 2\{[\mathrm{tr}(\mathbf{S})]^2 + \det(\mathbf{X} - \mathbf{X}^+)\det(\mathbf{B}_{22})\}^{1/2}.$$

3.4 Minimization of the object function for fixed values of $\mathbf{Y}_0$ or $\mathbf{R}_0$

For the fixed given value $\mathbf{Y}_0$ of the imaginary part of matrix $\mathbf{M}_0$, object function (55) is minimized by option (87), where $\mathbf{X}$ is given by (88).

For the fixed given value $\mathbf{R}_0$ of the real part of matrix $\mathbf{M}_0$, object function (55) is minimized by the imaginary part

(102) $$\mathbf{Y}_0 = \mathbf{B}_{22}^{-1/2}[\mathbf{S}^2 + \mathbf{B}_{22}^{1/2}(\mathbf{R}_0 - \mathbf{X})\mathbf{B}_{22}(\mathbf{R}_0 - \mathbf{X})^+\mathbf{B}_{22}^{1/2}]^{1/2}\mathbf{B}_{22}^{-1/2},$$

which is the result of (89).

4. IMPROVING THE NUMERICAL STABILITY OF THE OPTIMIZATION OF THE SHAPE OF GAUSSIAN BEAMS

4.1 Numerical quadrature

The real positive-definite symmetric 4 × 4 matrix for optimization (71), which is the result of the integration of the ordinary differential equations

$$\mathrm{d}\mathbf{B}/\mathrm{d}q_3 = \mathbf{\Pi}^+ \mathbf{F} \mathbf{\Pi}\,, \tag{103}$$

is ill-conditioned. It is usually not possible to compute matrix $\mathbf{C}_{11}$, see (90), composed of the submatrices (72) of $\mathbf{B}$, in single precision. This difficulty may be overcome by using the symmetric matrix

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{B}_{11} - \mathbf{B}_{12}\mathbf{B}_{22}^{-1}\mathbf{B}_{21} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix} \tag{104}$$

instead of $\mathbf{B}$ for numerical computations. We shall call matrix (104) the reduced matrix for optimization. In other words, the ill-conditioned matrix $\mathbf{B}$ may be expressed as the sum of two singular matrices

$$\mathbf{B} = \begin{pmatrix} \mathbf{C}_{12} \\ \mathbf{C}_{22} \end{pmatrix} \mathbf{C}_{22}^{-1}(\mathbf{C}_{21}, \mathbf{C}_{22}) + \begin{pmatrix} \mathbf{C}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}. \tag{105}$$

The differential equations for the reduced matrix $\mathbf{C}$ for optimization are

$$\begin{aligned} \mathrm{d}\mathbf{C}_{11}/\mathrm{d}q_3 &= (\mathbf{\Pi}_1 - \mathbf{\Pi}_2\mathbf{C}_{22}^{-1}\mathbf{C}_{21})^+ \,\mathbf{F}(\mathbf{\Pi}_1 - \mathbf{\Pi}_2\mathbf{C}_{22}^{-1}\mathbf{C}_{21})\,, \\ \mathrm{d}\mathbf{C}_{12}/\mathrm{d}q_3 &= \mathbf{\Pi}_1^+ \mathbf{F}\mathbf{\Pi}_2\,, \\ \mathrm{d}\mathbf{C}_{22}/\mathrm{d}q_3 &= \mathbf{\Pi}_2^+ \mathbf{F}\mathbf{\Pi}_2\,, \end{aligned} \tag{106}$$

where

$$\mathbf{\Pi}_1 = \begin{pmatrix} \mathbf{Q}_1 \\ \mathbf{P}_1 \end{pmatrix}, \quad \mathbf{\Pi}_2 = \begin{pmatrix} \mathbf{Q}_2 \\ \mathbf{P}_2 \end{pmatrix} \tag{107}$$

are 4 × 2 submatrices of the ray propagator matrix (10).

4.2 Operation of addition

We introduce the "addition" $\boldsymbol{+}$ of the reduced matrices $\mathbf{C}$ as

$$\mathbf{C}(\mathbf{F}^{(1)}) \boldsymbol{+} \mathbf{C}(\mathbf{F}^{(2)}) = \mathbf{C}(\mathbf{F}^{(1)} + \mathbf{F}^{(2)})\,, \tag{108}$$

see (58). Then the submatrices of the matrix

$$\mathbf{C} = \mathbf{C}^{(1)} \boldsymbol{+} \mathbf{C}^{(2)} \tag{109}$$

are given by

$$\begin{aligned} \mathbf{C}_{11} = \mathbf{C}_{11}^{(1)} + \mathbf{C}_{11}^{(2)} + &[\mathbf{C}_{12}^{(1)}(\mathbf{C}_{22}^{(1)})^{-1} - \mathbf{C}_{12}^{(2)}(\mathbf{C}_{22}^{(2)})^{-1}]\,. \\ &.\,[(\mathbf{C}_{22}^{(1)})^{-1} + (\mathbf{C}_{22}^{(2)})^{-1}]^{-1}\,. \\ &.\,[\mathbf{C}_{12}^{(1)}(\mathbf{C}_{22}^{(1)})^{-1} - \mathbf{C}_{12}^{(2)}(\mathbf{C}_{22}^{(2)})^{-1}]^+\,, \end{aligned}$$

$$\mathbf{C}_{12} = \mathbf{C}_{12}^{(1)} + \mathbf{C}_{12}^{(2)}, \tag{110}$$
$$\mathbf{C}_{22} = \mathbf{C}_{22}^{(1)} + \mathbf{C}_{22}^{(2)}.$$

4.3 Transformation of the reduced matrix for optimization

Assume that we have matrix (71) for optimization of the beam parameters at point $q_3 = q_3^{(0)}$ and that we need to know the matrix

$$\tilde{\mathbf{B}}(\mathbf{F}) = \int_{q_3(1)}^{q_3(2)} \mathbf{\Pi}^+(q_3, \tilde{q}_3^{(0)})\, \mathbf{F}(q_3)\, \mathbf{\Pi}(q_3, \tilde{q}_3^{(0)})\, \mathrm{d}q_3\,, \tag{111}$$

because we wish to optimize the beam parameters at point $q_3 = \tilde{q}_3^{(0)}$. Because of the chain property

$$\mathbf{\Pi}(q_3, \tilde{q}_3^{(0)}) = \mathbf{\Pi}(q_3, q_3^{(0)})\, \mathbf{\Pi}(q_3^{(0)}, \tilde{q}_3^{(0)}) \tag{112}$$

of the ray propagator matrix, equation (111) reads

$$\tilde{\mathbf{B}} = \tilde{\mathbf{\Pi}}^+ \mathbf{B} \tilde{\mathbf{\Pi}}\,, \tag{113}$$

where $\mathbf{B}$ is defined by (71) and

$$\tilde{\mathbf{\Pi}} = \mathbf{\Pi}(q_3^{(0)}, \tilde{q}_3^{(0)})\,. \tag{114}$$

As mentioned above, matrix $\mathbf{B}$ is ill-conditioned and we thus need to rewrite (113) in terms of the reduced matrix $\mathbf{C}$, see (104). Matrix $\mathbf{B}$ may be decomposed into two singular matrices, see (105),

$$\mathbf{B} = \begin{pmatrix} \mathbf{1} \\ \mathbf{0} \end{pmatrix} \mathbf{C}_{11}(\mathbf{1}, \mathbf{0}) + \begin{pmatrix} \mathbf{C}_{12}\mathbf{C}_{22}^{-1} \\ \mathbf{1} \end{pmatrix} \mathbf{C}_{22}(\mathbf{C}_{22}^{-1}\mathbf{C}_{21}, \mathbf{1})\,. \tag{115}$$

Equation (113) may then be rewritten as

$$\tilde{\mathbf{B}} = (\tilde{\mathbf{Q}}_1, \tilde{\mathbf{Q}}_2)^+\, \mathbf{C}_{11}(\tilde{\mathbf{Q}}_1, \tilde{\mathbf{Q}}_2) + (\mathbf{R}_1, \mathbf{R}_2)^+\, \mathbf{C}_{22}(\mathbf{R}_1, \mathbf{R}_2)\,, \tag{116}$$

where

$$(\tilde{\mathbf{Q}}_1, \tilde{\mathbf{Q}}_2) = (\mathbf{1}, \mathbf{0})\, \tilde{\mathbf{\Pi}}\,, \tag{117}$$

and

$$(\mathbf{R}_1, \mathbf{R}_2) = (\mathbf{C}_{22}^{-1}\mathbf{C}_{21}, \mathbf{1})\, \tilde{\mathbf{\Pi}}\,. \tag{118}$$

The reduced matrix (104) corresponding to $\tilde{\mathbf{B}}$ is

$$\tilde{\mathbf{C}} = \tilde{\mathbf{C}}^{(1)} + \tilde{\mathbf{C}}^{(2)}\,, \tag{119}$$

where

$$\tilde{\mathbf{C}}^{(1)} = \begin{pmatrix} \mathbf{0} & \tilde{\mathbf{Q}}_1^+ \mathbf{C}_{11} \tilde{\mathbf{Q}}_2 \\ \tilde{\mathbf{Q}}_2^+ \mathbf{C}_{11} \tilde{\mathbf{Q}}_1 & \tilde{\mathbf{Q}}_2^+ \mathbf{C}_{11} \tilde{\mathbf{Q}}_2 \end{pmatrix} \tag{120}$$

and

$$\tilde{\mathbf{C}}^{(2)} = \begin{pmatrix} \mathbf{0} & \mathbf{R}_1^+ \mathbf{C}_{22} \mathbf{R}_2 \\ \mathbf{R}_2^+ \mathbf{C}_{22} \mathbf{R}_1 & \mathbf{R}_2^+ \mathbf{C}_{22} \mathbf{R}_2 \end{pmatrix}. \tag{121}$$

Using expression (110), we arrive at

(122) $$\tilde{\mathbf{C}}_{11} = \mathbf{A}[(\tilde{\mathbf{C}}_{22}^{(1)})^{-1} + (\tilde{\mathbf{C}}_{22}^{(2)})^{-1}]^{-1}\,\mathbf{A}^+ ,$$

(123) $$\tilde{\mathbf{C}}_{12} = \tilde{\mathbf{C}}_{12}^{(1)} + \tilde{\mathbf{C}}_{12}^{(2)} ,$$

(124) $$\tilde{\mathbf{C}}_{22} = \tilde{\mathbf{C}}_{22}^{(1)} + \tilde{\mathbf{C}}_{22}^{(2)} ,$$

where

(125) $$\mathbf{A} = \tilde{\mathbf{Q}}_1^+(\tilde{\mathbf{Q}}_2^+)^{-1} - \mathbf{R}_1^+(\mathbf{R}_2^+)^{-1} ,$$

which is equivalent to

(126) $$\mathbf{A} = (\tilde{\mathbf{Q}}_2)^{-1}\,\tilde{\mathbf{Q}}_1 - \mathbf{R}_1^+(\mathbf{R}_2^+)^{-1}$$

and may be rearranged to read

(127) $$\mathbf{A} = (\tilde{\mathbf{Q}}_2)^{-1}(\tilde{\mathbf{Q}}_1, \tilde{\mathbf{Q}}_2)\,\Sigma(\mathbf{R}_1, \mathbf{R}_2)^+\,(\mathbf{R}_2^+)^{-1} .$$

Equations (117) and (118) may be inserted into (127) to yield

(128) $$\mathbf{A} = (\tilde{\mathbf{Q}}_2)^{-1}\,(\mathbf{1}, \mathbf{0})\,\tilde{\mathbf{\Pi}}\Sigma\tilde{\mathbf{\Pi}}^+ \begin{pmatrix} \mathbf{C}_{12}\mathbf{C}_{22}^{-1} \\ \mathbf{1} \end{pmatrix} (\mathbf{R}_2^+)^{-1} .$$

Considering that $\tilde{\mathbf{\Pi}}$ is a symplectic matrix, equation (128) simplifies to

(129) $$\mathbf{A} = (\tilde{\mathbf{Q}}_2)^{-1}\,(\mathbf{1}, \mathbf{0})\,\Sigma \begin{pmatrix} \mathbf{C}_{12}\mathbf{C}_{22}^{-1} \\ \mathbf{1} \end{pmatrix} (\mathbf{R}_2^+)^{-1}$$

and, after multiplication,

(130) $$\mathbf{A} = (\mathbf{R}_2^+\tilde{\mathbf{Q}}_2)^{-1} .$$

Equation (112) yields, for $q_3 = \tilde{q}_3^{(0)}$,

(131) $$\tilde{\mathbf{\Pi}} = \mathbf{\Pi}^{-1}(\tilde{q}_3^{(0)}, q_3^{(0)}) ,$$

and, since $\mathbf{\Pi}$ is a symplectic matrix,

(132) $$\tilde{\mathbf{\Pi}} = \Sigma^+\mathbf{\Pi}^+(\tilde{q}_3^{(0)}, q_3^{(0)})\,\Sigma .$$

We denote here the submatrices of the ray propagator matrix $\mathbf{\Pi}(\tilde{q}_3^{(0)}, q_3^{(0)})$ as

(133) $$\mathbf{\Pi}(\tilde{q}_3^{(0)}, q_3^{(0)}) = \begin{pmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \\ \mathbf{P}_1 & \mathbf{P}_2 \end{pmatrix} .$$

Then, see (132),

(134) $$\tilde{\mathbf{\Pi}} = \begin{pmatrix} \mathbf{P}_2^+ & -\mathbf{Q}_2^+ \\ -\mathbf{P}_1^+ & \mathbf{Q}_1^+ \end{pmatrix}$$

and equations (118), (120), (121), (122) read

(135) $$\mathbf{R}_1 = \mathbf{C}_{22}^{-1}\mathbf{C}_{21}\mathbf{P}_2^+ - \mathbf{P}_1^+ ,$$

(136) $$\mathbf{R}_2 = -\mathbf{C}_{22}^{-1}\mathbf{C}_{21}\mathbf{Q}_2^+ + \mathbf{Q}_1^+ ,$$

$$(137)\qquad \tilde{\mathbf{C}}_{12}^{(1)} = -\mathbf{P}_2\mathbf{C}_{11}\mathbf{Q}_2^+ , \quad \tilde{\mathbf{C}}_{12}^{(2)} = \mathbf{R}_1^+\mathbf{C}_{22}\mathbf{R}_2 ,$$

$$(138)\qquad \tilde{\mathbf{C}}_{22}^{(1)} = \mathbf{Q}_2\mathbf{C}_{11}\mathbf{Q}_2^+ , \quad \tilde{\mathbf{C}}_{22}^{(2)} = \mathbf{R}_2^+\mathbf{C}_{22}\mathbf{R}_2 ,$$

$$(139)\qquad \tilde{\mathbf{C}}_{11} = \{\mathbf{Q}_2\mathbf{R}_2[(\tilde{\mathbf{C}}_{22}^{(1)})^{-1} + (\tilde{\mathbf{C}}_{22}^{(2)})^{-1}]\,\mathbf{R}_2^+\mathbf{Q}_2^+\}^{-1} .$$

Equations (135) to (139) together with (123) and (124) may be used, e.g., for the transformation of the reduced matrix $\mathbf{C}$ for optimization from the initial point $q_3 = q_3^{(0)}$ to the endpoint $q_3 = \tilde{q}_3^{(0)}$ of a ray.

Acknowledgements: The principal part of this paper (except for Secs. 1 and 2.2) was written during the author's stay at the Institute of Geophysics of the Charles University in Prague. The author wishes to express his sincere thanks to Professor V. Červený and Dr. I. Pšenčík for valuable discussions and encouragement.

Received 12. 4. 1988 *Reviewer: I. Pšenčík*

References

[1] A. Ben-Menahem, W. B. Beydoun: Range of Validity of Seismic Ray and Beam Methods in General Inhomogeneous Media. Part I.: General Theory. Geophys. J. R. astr. Soc., 82 (1985), 207. Part II: A Canonical Problem. Geophys. J. R. astr. Soc., 82 (1985), 235.

[2] V. Červený: Gaussian Beam Synthetic Seismograms. J. Geophys., 58 (1985), 44.

[3] V. Červený, L. Klimeš, I. Pšenčík: Applications of the Dynamic Ray Tracing. Phys. Earth planet. Interiors, 51 (1988), 25.

[4] V. Farra, R. Madariaga: Seismic Waveform Modeling in Heterogeneous Media by Ray Perturbation Theory. J. geophys. Res., 92 (1987), No. B3, 2697.

[5] K. Fuchs, G. Müller: Computation of Synthetic Seismograms with the Reflectivity Method and Comparison with Observations. Geophys. J. R. astr. Soc. 23 (1971), 417.

[6] C. H. Chapman, R. Drummond: Body-Wave Seismograms in Inhomogeneous Media Using Maslov Asymptotic Theory. Bull. seismol. Soc. Am., 72 (1982), S277.

[7] A. P. Kachalov, M. M. Popov, I. Pšenčík: On the Application of the Method of Summation of Gaussian Beams to the Problems of Boundaries with Edges. In: V. M. Babich (ed.): Mathematical Problems of the Theory of Propagation of Waves, Vol. 13, Nauka, Leningrad 1983, pp. 65—71 (in Russian).

[8] W. Kim, J. Garmany: Optimum Beamwidth for Gaussian Beams in Three Dimensional Problems. EOS, 66 (1985), 980.

[9] K. D. Klem-Musatov, A. M. Aizenberg: The Ray Method and the Theory of Edge Waves. Geophys. J. R. astr. Soc., 79 (1984), 35.

[10] 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.

[11] J. Konopásková, V. Červený: Numerical Modelling of Time-Harmonic Seismic Wave Fields in Simple Structures. Part I: Stud. geophys. geod., 28 (1984), 19. Part II: Stud. geophys. geod., 28 (1984), 113.

[12] R. Madariaga: Gaussian Beam Synthetic Seismograms in a Vertically Varying Medium. Geophys. J. R. astr. Soc., 79 (1984), 589.

[13] M. Musil: Seismic Wavefields in Three-Dimensional Absorbing Media (With Application to Vertical Seismic Profiling). Diploma thesis, Faculty of Mathematics and Physics, Charles University, Prague 1986 (in Czech, not published).

[14] R. Nowack, K. Aki: The Two-Dimensional Gaussian Beam Synthetic Method: Testing and Application. J. geophys. Res., 89 (1984), No. B9, 7797.

[15] C. J. Thomson, C. H. Chapman: An Introduction to Maslov's Asymptotic Method. Geophys. J. R. astr. Soc., 83 (1985), 143.

[16] M. Weber: Computation of Body-Wave Seismograms in Absorbing 2-D Media Using the Gaussian Beam Method: Comparison with Exact Methods. Geophys. J., 92 (1988), 9.

[17] 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.