Wave Inversion Technology, Report No. 3, pages 143-155

Three-Dimensional Dynamic Wavefield Properties

C. Vanelle and D. Gajewski[1]

keywords: *ray propagator, paraxial rays, traveltimes, geometrical spreading, amplitudes*

ABSTRACT

Information about dynamic wavefield properties can be expressed in terms of the ray propagator which depends on second order derivatives of traveltimes. Research concerning this topic has already been carried out by various groups. It was, however, restricted to surfaces either defined by reflectors or the ray centered coordinate system. In this paper we present a technique that uses the three-dimensional cartesian grid defined by traveltime data tables. It is particularly suited for pre-stack migration applications since there the data is available in this format. A fully 3-D 6$\times$6 ray propagator matrix is introduced. It contains all information needed to compute the required properties and can be determined from traveltimes only. Its relation to the commonly used 4$\times$4 ray propagator matrices is given. Applications shown to demonstrate the versatility of the new method are traveltime interpolations and the computation of geometrical spreading.

INTRODUCTION

The foundation for a 3-D pre-stack migration of the Kirchhoff type is a summation stack along diffraction surfaces. To accomplish this accurately and efficiently is still a challenge. Proper migration weights have to be applied to achieve high accuracy. To keep the computational costs within reasonable limits it is important to optimize the migration aperture which depends on information about the Fresnel zones. To reach both objectives requires knowledge about dynamic wavefield properties. Although a number of extremely efficient tools for traveltime computation have been developed in recent years (e.g., Ettrich and Gajewski (1996), Leidenfrost (1998)), fast algorithms like FD techniques yield only kinematic but no dynamic informations. Publications by Bortfeld (1989) and Hubral et al. (1992) make use of the relationship between the curvature of a wavefront and the ray propagator matrix to determine dynamic wavefield properties. Their approach is, however, confined to the reference surfaces they work

[1]**email:** vanelle@dkrz.de

on, e.g., the wavefront or reflector surfaces, whereas seismic traveltime data is most likely to be sampled or computed on rectangular grids that do not coincide with these surfaces. Therefore, the application of their algorithms is rather inconvenient. In the following we will show how a newly introduced 6×6 ray propagator matrix can be applied to circumvent the restriction to surfaces and thus make it possible to work directly on the (cartesian) traveltime grids.

METHOD

Traveltime Expansions

We begin with the assumption of a smooth traveltime field. In real life this is justified by the Earth's low pass filter behaviour. From the computational point of view it means that velocity variations must not take place on a smaller scale than the signal's wavelength. This condition corresponds to the applicability of the ray method. With this requirement fulfilled, we can approximate the traveltime difference $\tau(\hat{\vec{s}}', \hat{\vec{g}}')$ between a point $\hat{\vec{s}}'$ and a point $\hat{\vec{g}}'$, see Figure 1, by a Taylor expansion. For small $\hat{\vec{s}}$ and $\hat{\vec{g}}$ we get the following expression for $\tau(\hat{\vec{s}}', \hat{\vec{g}}')$:

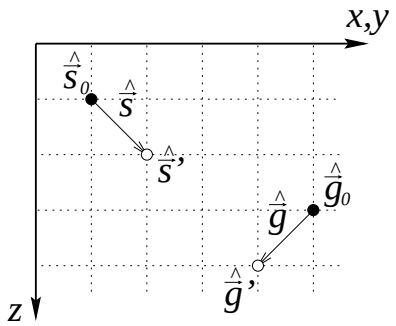

Figure 1: Nomenclature used: we denote the deviation in source position by $\hat{\vec{s}} = \hat{\vec{s}}' - \hat{\vec{s}}_0$ and in geophone position by $\hat{\vec{g}} = \hat{\vec{g}}' - \hat{\vec{g}}_0$. The hat on top of a vector distinguishes between a two-dimensional vector $\vec{v}$ and a three-dimensional vector $\hat{\vec{v}}$. The same notation is applied to matrices.

$$\tau(\hat{\vec{s}}', \hat{\vec{g}}') = \tau_0 - \hat{\vec{p}}\,\hat{\vec{s}} + \hat{\vec{q}}\,\hat{\vec{g}} - \frac{1}{2}\,\hat{\vec{s}}^{\,T}\hat{\mathcal{S}}\,\hat{\vec{s}} + \frac{1}{2}\,\hat{\vec{g}}^{\,T}\hat{\mathcal{G}}\,\hat{\vec{g}} - \hat{\vec{s}}^{\,T}\hat{\mathcal{N}}\,\hat{\vec{g}} + O(3). \qquad (1)$$

The first order derivatives $-\hat{\vec{p}}$ and $\hat{\vec{q}}$ are the slowness vectors at source and receiver and τ_0 is the traveltime difference between two points $\hat{\vec{s}}_0$ and $\hat{\vec{g}}_0$. The coefficient matrices $\hat{\mathcal{S}}$, $\hat{\mathcal{G}}$ and $\hat{\mathcal{N}}$ are

$$\hat{\mathcal{S}}_{ij} = -\frac{\partial^2\tau}{\partial s_i\,\partial s_j}, \quad \hat{\mathcal{G}}_{ij} = \frac{\partial^2\tau}{\partial g_i\,\partial g_j}, \quad \hat{\mathcal{N}}_{ij} = -\frac{\partial^2\tau}{\partial s_i\,\partial g_j} \qquad (2)$$

The signs in (1) take into account that $\tau(\hat{\vec{s}}', \hat{\vec{g}}')$ is the traveltime *difference* between the two points. We make a further approximation: as seismic traveltimes are better

approximated by hyperbolae than by parabolae (e.g., Ursin (1982)) we mold (1) into a hyperbolic form as shown by Schleicher et al. (1993). We square (1) and neglect terms of higher order than two, resulting in

$$\tau^2(\hat{\vec{s}}\,',\hat{\vec{g}}\,') = (\tau_0 - \hat{\vec{p}}\,\hat{\vec{s}} + \hat{\vec{q}}\,\hat{\vec{g}})^2 + \tau_0(-\hat{\vec{s}}^{\,T}\hat{\mathcal{S}}\,\hat{\vec{s}} + \hat{\vec{g}}^{\,T}\hat{\mathcal{G}}\,\hat{\vec{g}} - 2\,\hat{\vec{s}}^{\,T}\hat{\mathcal{N}}\,\hat{\vec{g}}) \;+\; O(3). \qquad (3)$$

The ability to approximate traveltimes by (1) or (3) can be inverted: it also means that if traveltime information is available for various source and receiver combinations as, e.g., from multi-coverage experiments, then equations (1) and (3) can be used for determining the coefficients by solving for them. This was initially suggested by Bortfeld (1989) and first explained by Gajewski (1998). It not only offers the possibility to compute traveltimes for intermediate source and receiver positions but can also lead to an immense saving in memory: the valid vicinity of the approximations is not neccessarily restricted to adjoining gridpoints of a discretized subsurface model. Therefore we can revert to storing only a fraction of the traveltime information on a coarse grid and reconstruct the remaining data by interpolation using (1) or (3). The degree of redundancy in the initial traveltime information depends on the model under consideration. To give an example: if we use only every tenth gridpoint in three dimensions we need a factor 10^3 less in storage capacity! Please note also that the method is not restricted to cubical grids.

Geometrical Spreading

We will now compare equation (1) to the paraxial traveltime equation by Cervený (1987):

$$\tau(\hat{\vec{s}}_r\,',\hat{\vec{g}}_r\,') = \tau_0 - \hat{\vec{p}}_r\,\hat{\vec{s}}_r + \hat{\vec{q}}_r\,\hat{\vec{g}}_r - \frac{1}{2}\,\hat{\vec{s}}_r^{\,T}\hat{\mathcal{M}}_S\,\hat{\vec{s}}_r + \frac{1}{2}\,\hat{\vec{g}}_r^{\,T}\hat{\mathcal{M}}_G\,\hat{\vec{g}}_r - \vec{s}_r^{\,T}\mathcal{Q}_2^{-1}\,\vec{g}_r. \qquad (4)$$

The index r denotes ray centered coordinates. Note that there are no third components in the last term. The transformation between the ray centered and the cartesian system is a rotation employing two angles ϑ and φ. The matrices $\hat{\mathcal{M}}$ and $\mathcal{Q}_2$ used by Cervený and the transformation between ray centered and cartesian system are explained in the appendix. Cervený (1987) shows that the modulus of the relative geometrical spreading $\mathcal{L}$ depends only on the determinant of $\mathcal{Q}_2$. With v_s being the velocity at the source it reads

$$|\mathcal{L}| = \frac{1}{v_s}\sqrt{|\mathcal{Q}_2|}. \qquad (5)$$

In the appendix we show that there are only four independant components to matrix $\hat{\mathcal{N}}$ and how $\hat{\mathcal{N}}$ relates to $\mathcal{Q}_2^{-1}$. From that we can compute $|\mathcal{Q}_2|$ by the determinant of any 2×2 submatrix of $\hat{\mathcal{N}}$ and the appropriate angles, e.g., for the upper left submatrix of $\hat{\mathcal{N}}$ we find

$$|\mathcal{L}| = \frac{1}{v_s}\sqrt{\frac{\cos\vartheta_s\cos\vartheta_g}{|N_{xx}N_{yy} - N_{xy}N_{yx}|}}. \qquad (6)$$

Hubral et al. (1992) do a similar expansion as (1) into an anterior and a posterior surface using $\vec{s}$, $\vec{g}$ instead of $\hat{\vec{s}}$, $\hat{\vec{g}}$. If their surfaces equal the z-plane we use, their result corresponds to equation (6). In this case the ϑ are the incidence respectively emergence angles.

APPLICATIONS

Hyperbolic vs. Parabolic Expansion

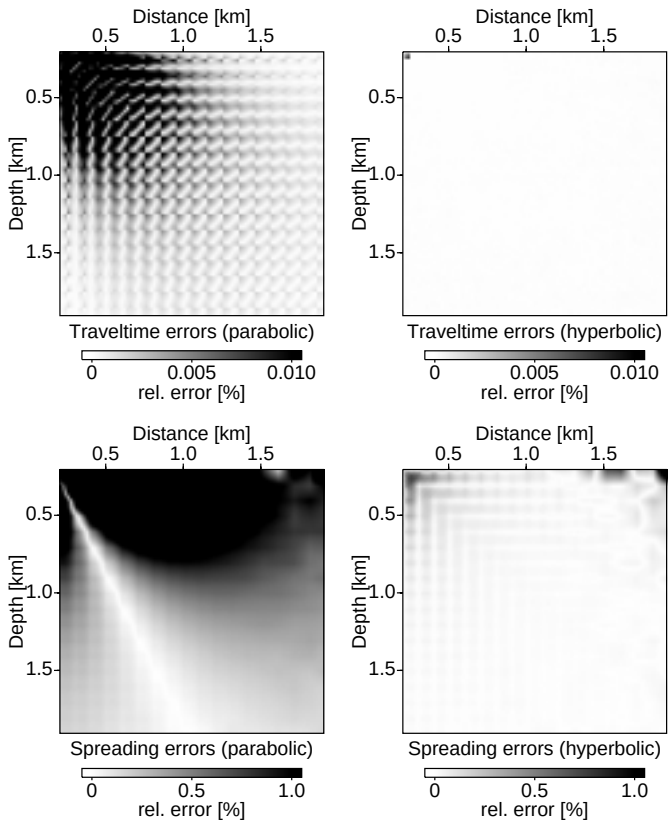

Figure 2: Comparison of parabolic (left) and hyperbolic (right) expansion for a constant velocity model. For both traveltimes (top) and geometrical spreading (bottom) the hyperbolic variant is far superior.

Knowing that diffraction traveltimes can be approximated by hyperbolae rather than by parabolae we expected better results for the hyperbolic variant. This was confirmed by two-dimensional examples in traveltime interpolation and computation of geometrical spreading. We used a constant velocity model with 100m gridspacing to compute the coefficients of (1) and (3). For the traveltime interpolation onto the fine grid with 10m spacing (1) respectively (3) were applied. For the computation of

geometrical spreading (6) was used on the coarse grid. Bilinear interpolation was then carried out for the intermediate points on the fine grid. The resulting errors are shown in Figure 2. For the traveltimes the hyperbolic variant has an average error within machine precision. i.e., 10^{-5}% (maximum 0.01%) whereas the parabolic variant's average error is 0.01% (3.1% max.). Note that a bilinear traveltime interpolation would yield 0.6% (>8% max.). For the geometrical spreading we get an average error of 0.03% (2.3% max.) for the hyperbolic and 1.0% (34.3% max.) for the parabolic variant. Due to these results we will not use the parabolic approximation any further.

Traveltime Interpolation

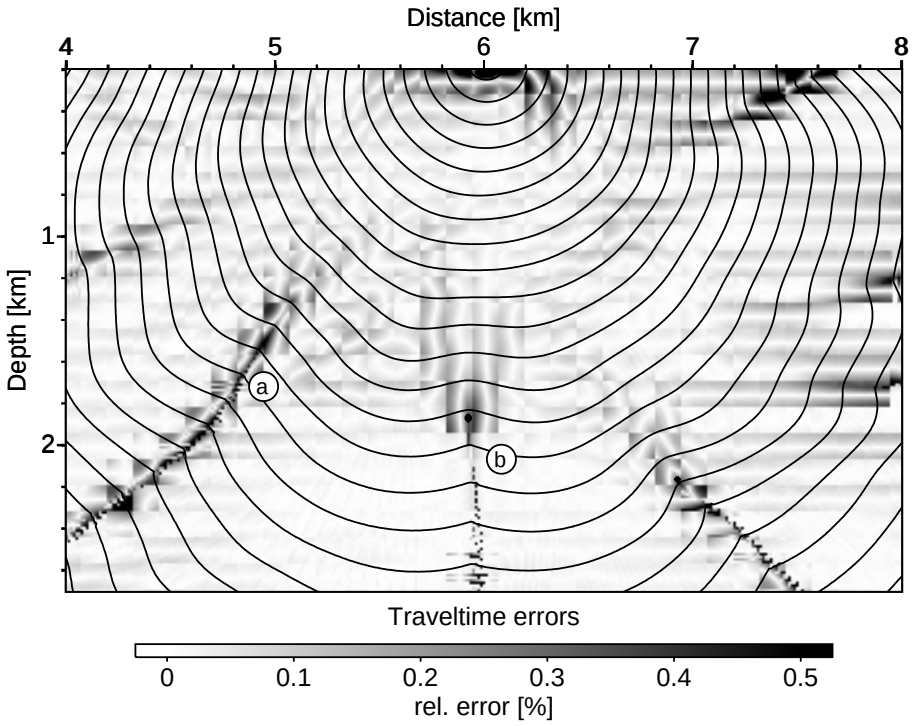

Figure 3: Relative traveltime errors and isochrones for the Marmousi model. Some errors occur in the vicinity of edges in the wavefront, see the area indicated with ⓐ. The reason is that a Taylor expansion is not valid there because the traveltimes are not smooth. It is, however, possible to locate these regions and perform a single-sided extrapolation. This was done for the triplication starting at ⓑ for all later times.

We give two more examples for traveltime interpolation. The first is a two dimensional version of the Marmousi model described by Versteeg and Grau (1991) with 125m (coarse) grid spacing in either direction. As traveltimes were computed by wave front construction the model was 100 fold smoothed. Interpolation onto a 12.5m

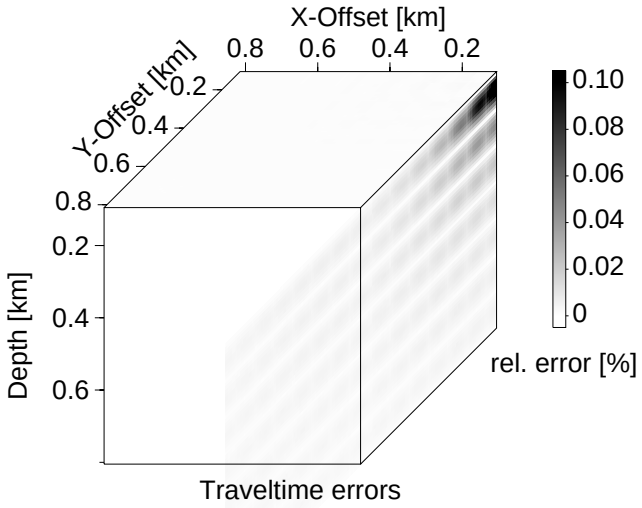

Figure 4: Relative traveltime errors for a constant velocity gradient model. The area on top of the cube has indeed as low an error as the white colour indicates.

spaced fine grid yields a median error of only 0.02%. A second example is a three dimensional constant velocity gradient model with 100m coarse grid spacing and a 10m fine grid. Here we find the median error to be 0.002%.

Geometrical Spreading

We show results for the same two models as in the traveltime section. Bilinear interpolation was applied for the computation onto the fine grid. For the Marmousi model we used a 62.5m coarse and a 12.5m fine grid. The median error is 1.8%. The dominating errors occur in the vicinity of edges in the wavefront (Figure 5) where different phases get mixed if we use first arrivals only. Employing later arrivals and separately interpolating for both branches of the traveltime curve would reduce the error in these regions. The second example is again a three-dimensional constant velocity gradient model. The same grid spacings as for the traveltimes are used. Figure 6 finds the median error to be smaller than 0.3%. Generally the error increases for higher offsets and for too small coarse grid spacing as the moveout gets smaller and traveltime inaccuracies have a higher impact.

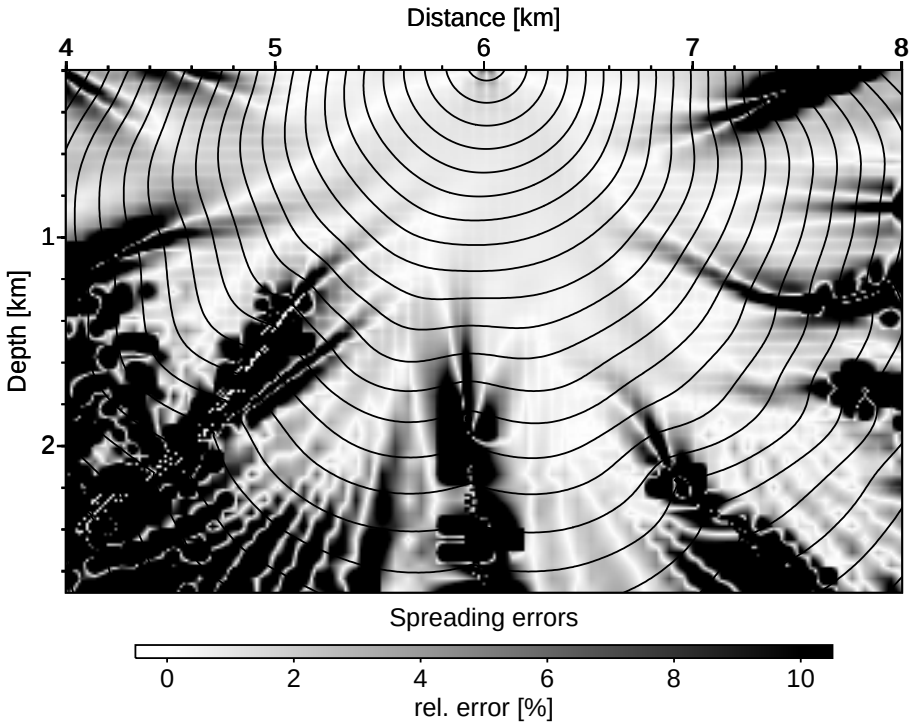

Figure 5: Relative errors in geometrical spreading and isochrones for the Marmousi model. Again, as for the traveltimes, the interpolation fails in regions where triplications occur. Due to the bilinear interpolation the effect has a stronger impact here.

CONCLUSIONS AND DISCUSSION

A new technique for the determination of a three-dimensional ray propagator matrix $\underline{\underline{\hat{T}}}$ was presented. The theory links this matrix to the established propagator matrices $\underline{\underline{\Pi}}$ (Cervený's formulation) and $\underline{\underline{T}}$ (by Bortfeld and Hubral). Two applications were shown to be very effective and accurate. Hyperbolic interpolation of traveltimes leads to immense savings in storage at no significant loss in accuracy. Geometrical spreading can be computed from traveltimes only. These results make further applications look highly promising. There is only a small step to be taken from geometrical spreading to the computation of migration weights. The determination of Fresnel zones is also possible with the ray propagator. Thus we have the means to provide an efficient algorithm for true amplitude migration with no significant additional effort since the required traveltime tables are needed for Kirchhoff-type pre-stack migration. This algorithm will be memory efficient since every computation can be carried out with coarse grid traveltime data only and time efficient as we can optimize the migration aperture by knowledge about the Fresnel zones.

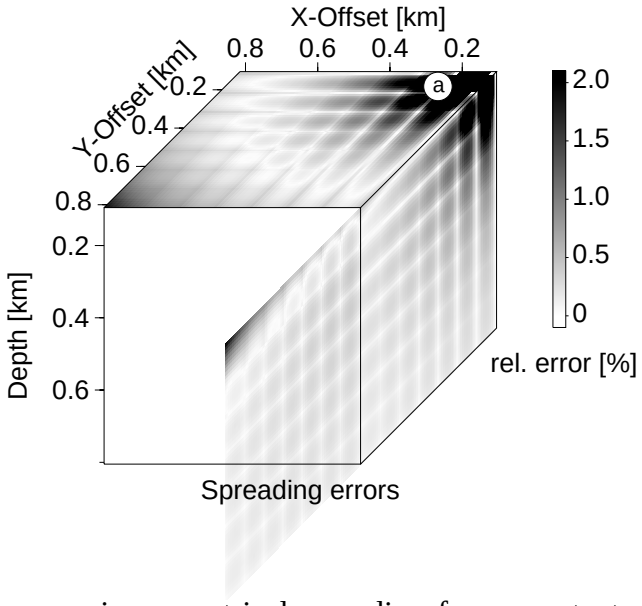

Figure 6: Relative errors in geometrical spreading for a constant velocity gradient model. Only near the source we have higher errors, see ⓐ. The reason is that the radius of the wavefront curvature is small compared to the coarse grid spacing: bilinear interpolation does not fit well. This is not problematic since we are not interested in the source region.

ACKNOWLEDGEMENTS

This work was partially supported by the European Commission (JOF3-CT97-0029) and the sponsors of the WIT consortium. Continous discussions with the members of the Applied Geophysics Group Hamburg are appreciated, especially with Boris Kashtan.

REFERENCES

Bortfeld, R., 1989, Geometrical ray theory: Rays and traveltimes in seismic systems (second-order approximation of traveltimes): Geophysics, **54**, 342–349.

Cervený, V., 1987, Ray methods for three-dimensional seismic modeling: Norwegian Institute for Technology (Petroleum Industry Course).

Ettrich, N., and Gajewski, D., 1996, Wave front construction in smooth media for prestack depth migration: PAGEOPH, **148**, 481–502.

Gajewski, D., and Vanelle, C., 1999, Computing wavefront characteristics without

dynamic ray tracing: 61th Ann. Internat. Mtg., Eur. Assn. Expl. Geophys., Expanded Abstracts, 4.14.

Gajewski, D., 1998, Determining the ray propagator from traveltimes: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1900–1903.

Hubral, P., Schleicher, J., and Tygel, M., 1992, Three-dimensional paraxial ray properties, part 1: Basic relations: J. Seis. Expl., **1**, 265–279.

Leidenfrost, A., 1998, Fast computation of traveltimes in two and three dimensions: Ph.D. thesis, University of Hamburg.

Schleicher, J., Tygel, M., and Hubral, P., 1993, Parabolic and hyperbolic paraxial two-point traveltimes in 3d media: Geophysical Prospecting, **41**, 495–513.

Ursin, B., 1982, Quadratic wavefront and traveltime approximations in inhomogeneous layered media with curved interfaces: Geophysics, **47**, 1012–1021.

Vanelle, C., and Gajewski, D., 1999, Determining geometrical spreading from traveltimes: 61th Ann. Internat. Mtg., Eur. Assn. Expl. Geophys., Expanded Abstracts, 1.56.

Versteeg, R., and Grau, G., 1991, The Marmousi experience: 1990 EAEG workshop on practical aspects of seismic data inversion, Zeist, Netherlands, Proceedings.

PUBLICATIONS

Previous results concerning the determination of the Bortfeld ray propagator were published by Gajewski and Vanelle (1999) and Gajewski (1998). Amplitude considerations were presented by Vanelle and Gajewski (1999).

APPENDIX A

We introduce a matrix $\underline{\underline{\hat{T}}}$ for describing the propagation of a ray in a three dimensional medium in cartesian coordinates. A similar propagator matrix $\underline{\underline{T}}$ was introduced by Bortfeld (1989) for projections of the slownesses and coordinates onto reference surfaces. Our propagator matrix however is not restricted to surfaces but uses third components as well. We use Hamiltons equation (Bortfeld (1989))

$$\mathrm{d}\tau = \hat{\vec{q}}'\mathrm{d}\hat{\vec{g}} - \hat{\vec{p}}'\mathrm{d}\hat{\vec{s}} \qquad \text{(A-1)}$$

and assume a linear relationship between the source and receiver variables as Hubral et al. (1992) do, but for three dimensions:

$$\begin{aligned} \hat{\vec{g}} &= \hat{\mathcal{A}}\,\hat{\vec{s}} + \hat{\mathcal{B}}\,\hat{\vec{p}}\,' - \hat{\vec{p}} \\ \hat{\vec{q}}\,' - \hat{\vec{q}} &= \hat{\mathcal{C}}\,\hat{\vec{s}} + \hat{\mathcal{D}}\,\hat{\vec{p}}\,' - \hat{\vec{p}} \end{aligned} \qquad \text{(A-2)}$$

or rewritten as

$$\begin{pmatrix} \hat{\vec{g}} \\ \hat{\vec{q}}\,' - \hat{\vec{q}} \end{pmatrix} = \underline{\underline{\hat{T}}} \begin{pmatrix} \hat{\vec{s}} \\ \hat{\vec{p}}\,' - \hat{\vec{p}} \end{pmatrix}, \quad \text{with} \quad \underline{\underline{\hat{T}}} = \begin{pmatrix} \hat{\mathcal{A}} & \hat{\mathcal{B}} \\ \hat{\mathcal{C}} & \hat{\mathcal{D}} \end{pmatrix}. \qquad \text{(A-3)}$$

We solve (A-2) for $\hat{\vec{q}}\,'$ and $\hat{\vec{p}}\,'$ and insert the result into (A-1), yielding

$$\mathrm{d}\tau = (\hat{\vec{q}} + \hat{\mathcal{D}}\,\hat{\mathcal{B}}^{-1}\hat{\vec{g}} - \hat{\mathcal{D}}\,\hat{\mathcal{B}}^{-1}\hat{\mathcal{A}}\,\hat{\vec{s}} + \hat{\mathcal{C}}\,\hat{\vec{s}})\,\mathrm{d}\hat{\vec{g}} + (-\hat{\vec{p}} - \hat{\mathcal{B}}^{-1}\hat{\vec{g}} + \hat{\mathcal{B}}^{-1}\hat{\mathcal{A}}\,\hat{\vec{s}})\,\mathrm{d}\hat{\vec{s}}. \qquad \text{(A-4)}$$

This expression can only be integrated if $\mathrm{d}\tau$ is a total differential. This condition leads to the following relationships, cf. Hubral et al. (1992):

$$\hat{\mathcal{A}}^T\hat{\mathcal{D}} - \hat{\mathcal{C}}^T\hat{\mathcal{B}} = \underline{\hat{1}} \qquad \hat{\mathcal{D}}^T\hat{\mathcal{B}} = \hat{\mathcal{B}}^T\hat{\mathcal{D}} \qquad \hat{\mathcal{A}}^T\hat{\mathcal{C}} = \hat{\mathcal{C}}^T\hat{\mathcal{A}}, \qquad \text{(A-5)}$$

where $\underline{\hat{1}}$ is the identity matrix. The resulting traveltime difference is

$$\tau(\hat{\vec{s}}\,', \hat{\vec{g}}\,') = \tau_0 - \hat{\vec{p}}\,\hat{\vec{s}} + \hat{\vec{q}}\,\hat{\vec{g}} - \frac{1}{2}\,\hat{\vec{s}}^T\hat{\mathcal{D}}\,\hat{\mathcal{B}}^{-1}\,\hat{\vec{s}} + \frac{1}{2}\,\hat{\vec{g}}^T\hat{\mathcal{B}}^{-1}\hat{\mathcal{A}}\,\hat{\vec{g}} - \hat{\vec{s}}^T\hat{\mathcal{B}}^{-1}\,\hat{\vec{g}} \qquad \text{(A-6)}$$

where τ_0 is the traveltime from $\hat{\vec{s}}_0$ to $\hat{\vec{g}}_0$. If we compare equations (1) and (A-6) and use (A-5) we find that

$$\hat{\mathcal{S}} = \hat{\mathcal{S}}^T = -\hat{\mathcal{B}}^{-1}\hat{\mathcal{A}} \qquad \hat{\mathcal{G}} = \hat{\mathcal{G}}^T = \hat{\mathcal{D}}\,\hat{\mathcal{B}}^{-1} \qquad \hat{\mathcal{N}} = \hat{\mathcal{B}}^{-1} \neq \hat{\mathcal{N}}^T \qquad \text{(A-7)}$$

and this results in the ingredients of $\underline{\underline{\hat{T}}}$:

$$\underline{\underline{\hat{T}}} = \begin{pmatrix} -\hat{\mathcal{N}}^{-1}\hat{\mathcal{S}} & \hat{\mathcal{N}}^{-1} \\ -\hat{\mathcal{N}}^T - \hat{\mathcal{G}}\,\hat{\mathcal{N}}^{-1}\hat{\mathcal{S}} & \hat{\mathcal{G}}\,\hat{\mathcal{N}}^{-1} \end{pmatrix}. \qquad \text{(A-8)}$$

APPENDIX B

Cervený (1987) introduces the ray propagator matrix $\underline{\underline{\mathrm{II}}}$ with the components

$$\underline{\underline{\mathrm{II}}} = \begin{pmatrix} \mathcal{Q}_1 & \mathcal{Q}_2 \\ \mathcal{P}_1 & \mathcal{P}_2 \end{pmatrix} \qquad \text{(B-1)}$$

where the $\mathcal{Q}_i$, $\mathcal{P}_i$ are 2×2 matrices. Equation (4) contains them in the following form:

$$\hat{\mathcal{M}}_s = \begin{pmatrix} -\mathcal{Q}_2^{-1}\mathcal{Q}_1 & \begin{matrix} -\frac{1}{v_s^2}\frac{\partial v_s}{\partial s_{r_x}} \\ -\frac{1}{v_s^2}\frac{\partial v_s}{\partial s_{r_y}} \end{matrix} \\ \begin{matrix} -\frac{1}{v_s^2}\frac{\partial v_s}{\partial s_{r_x}} & -\frac{1}{v_s^2}\frac{\partial v_s}{\partial s_{r_y}} \end{matrix} & -\frac{1}{v_s^2}\frac{\partial v_s}{\partial s_{r_z}} \end{pmatrix} \qquad \text{(B-2)}$$

and

$$\hat{\mathcal{M}}_g = \begin{pmatrix} & & -\frac{1}{v_g^2}\frac{\partial v_g}{\partial g_{rx}} \\ \mathcal{P}_2\mathcal{Q}_2^{-1} & & -\frac{1}{v_g^2}\frac{\partial v_g}{\partial g_{ry}} \\ -\frac{1}{v_g^2}\frac{\partial v_g}{\partial g_{rx}} & -\frac{1}{v_g^2}\frac{\partial v_g}{\partial g_{ry}} & -\frac{1}{v_g^2}\frac{\partial v_g}{\partial g_{rz}} \end{pmatrix}. \tag{B-3}$$

The remaining matrix $\mathcal{P}_1$ is connected with $\mathcal{P}_2$, $\mathcal{Q}_1$, $\mathcal{Q}_2$ by symplecticity relations similar to (A-5), cf. Cervený (1987). Important properties like geometrical spreading can be computed from the ray propagator in ray centered coordinates. We will now show that with a transformation from ray centered to our cartesian coordinates we can derive similar relationships for our cartesian propagator. The difference between (4) and (1) lies in the coordinate systems. Both use two cartesian systems centered in $\hat{\vec{s}}_0 = \hat{\vec{s}}_{r_0}$ ($\hat{\vec{g}}_0 = \hat{\vec{g}}_{r_0}$). The transformation between them is given by a rotation matrix $\hat{\mathcal{R}}_s$ ($\hat{\mathcal{R}}_g$) such that

$$\hat{\vec{s}} = \hat{\mathcal{R}}_s \hat{\vec{s}}_r \qquad (\hat{\vec{g}} = \hat{\mathcal{R}}_g \hat{\vec{g}}_r). \tag{B-4}$$

Our choice of $\hat{\mathcal{R}}_s$ ($\hat{\mathcal{R}}_g$) is that the ray centered system is first rotated by an angle ϑ_s (ϑ_g) until the 3-components coincide. Then a second rotation around x_3 by angle φ_s (φ_g) adjusts the remaining 1- and 2-components. The matrix $\hat{\mathcal{R}}_s$ looks as follows:

$$\hat{\mathcal{R}}_s = \begin{pmatrix} \cos\vartheta_s\cos\varphi_s & \sin\varphi_s & -\sin\vartheta_s\cos\varphi_s \\ -\cos\vartheta_s\sin\varphi_s & \cos\varphi_s & \sin\vartheta_s\cos\varphi_s \\ \sin\vartheta_s & 0 & \cos\vartheta_s \end{pmatrix}. \tag{B-5}$$

The angles can be easily obtained from the cartesian slowness vector

$$\hat{\vec{p}} = (p_x, p_y, p_z) = \hat{\vec{p}}_r \hat{\mathcal{R}}_S^T \tag{B-6}$$

and $\hat{\vec{p}}_r = (0, 0, 1/v_s^2)$. The matrices $\hat{\mathcal{S}}$ ($\hat{\mathcal{G}}$) transform

$$\hat{\mathcal{M}}_s = \hat{\mathcal{R}}_s^T \hat{\mathcal{S}} \hat{\mathcal{R}}_s \qquad (\hat{\mathcal{M}}_g = \hat{\mathcal{R}}_g^T \hat{\mathcal{G}} \hat{\mathcal{R}}_g). \tag{B-7}$$

For the mixed derivative matrix $\hat{\mathcal{N}}$ we find that there are only four independant components as, e.g.,

$$\begin{aligned} N_{zx} &= \frac{\partial p_z}{\partial x_g} = \frac{\partial}{\partial x_g}\sqrt{\frac{1}{v_s^2} - p_x^2 - p_y^2} = -\frac{1}{p_z}\left(p_x\frac{\partial p_x}{\partial x_g} + p_y\frac{\partial p_y}{\partial x_g}\right) \\ &= \tan\vartheta_s(N_{xx}\cos\varphi_s - N_{yx}\sin\varphi_s). \end{aligned} \tag{B-8}$$

Expressions for the other components can be found in a similar way. If we assume a matrix $\hat{\mathcal{Q}}$ to be a three-dimensional expansion of $\mathcal{Q}_2^{-1}$ with

$$\hat{\mathcal{Q}} = \begin{pmatrix} & & 0 \\ \mathcal{Q}_2^{-1} & & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{B-9}$$

the transformation

$$\hat{\mathcal{Q}} = \hat{\mathcal{R}}_s^T \hat{\mathcal{N}} \hat{\mathcal{R}}_g \tag{B-10}$$

yields indeed disappearing 3-components of $\hat{\mathcal{Q}}$ if we use (B-8) etc., thus

$$\hat{\vec{s}}^T \hat{\mathcal{N}} \hat{\vec{g}} = \hat{\vec{s}}_r{}^T \hat{\mathcal{Q}} \hat{\vec{g}}_r = \vec{s}_r{}^T \mathcal{Q}_2^{-1} \vec{g}_r. \tag{B-11}$$

This result means that $\mathcal{Q}_2^{-1}$ can be expressed by any 2×2 submatrix of $\hat{\mathcal{N}}$ and the corresponding rotation angles.