Geophys. J. Int. (2002) **150,** 440–466

Fréchet kernels for body-wave amplitudes

F. A. Dahlen and Adam M. Baig

Department of Geosciences, Princeton University, Princeton, NJ 08544, *USA. E-mails: fad@princeton.edu; abaig@princeton.edu*

Accepted 2002 February 13. Received 2002 January 22; in original form 2001 June 22

SUMMARY

We derive a 3D Fréchet sensitivity kernel relating the rms amplitude of a far-field, broad-band body-wave pulse to laterally heterogeneous seismic slowness variations within the earth. Unlike the 'banana–doughnut' sensitivity kernel for a cross-correlation traveltime, the amplitude Fréchet kernel for a turning wave is maximally sensitive, rather than completely insensitive, to the 3D slowness perturbation along the central source-to-receiver ray. In the asymptotic limit of an infinite-frequency pulse, our 3D amplitude kernel formulation is consistent with the dominant 1D integral involving the cross-path curvature of the slowness perturbation along the unperturbed geometrical ray.

Key words: body waves, Fréchet derivatives, global seismology, perturbation methods, ray theory, tomography.

1 INTRODUCTION

The amplitude of a teleseismic body-wave pulse is sensitive to 3D seismic slowness variations within the earth, as a result of geometrical focusing and defocusing and finite-frequency diffraction and interference effects. Because of the difficulty of disentangling these purely elastic effects from 3D variations in anelastic attenuation and moment-tensor errors, it will not be an easy matter to assimilate body-wave amplitude measurements into whole-earth tomographic inversion studies. A necessary first step, however, is the development of a 3D Fréchet sensitivity kernel which relates a finite-frequency amplitude measurement to laterally heterogeneous variations in elastic slowness. We use the Born approximation to derive such an elastic amplitude Fréchet kernel here.

Although we have striven for succinctness, it is an unfortunate fact that this paper is replete with long-winded mathematical manipulations; for this reason, a guided tour of the contents may be helpful. The 3D slowness sensitivity kernel itself is actually fairly easy to derive; the analysis is given in Section 2, and some illustrative examples of both traveltime and amplitude Fréchet kernels for selected teleseismic phases are depicted in Section 4. The remainder of the paper is concerned with an important but somewhat ancillary theoretical question—*does the 3D amplitude kernel formulation reduce to the corresponding result obtained using geometrical ray perturbation theory in the limit of infinite frequency*, as expected? The asymptotic reduction of the 3D Fréchet integral to a 1D integral along the central geometrical ray in the limit $\omega \to \infty$ is described in Section 3. In Sections 5 and 6, which together account for more than half the total number of paragraphs and equations, we present an independent *ab initio* analysis of the ray-theoretical amplitude perturbation, based upon a Hamiltonian formulation of the paraxial ray-tracing equations. Readers who are not interested in the technical details of asymptotic 3D integral evaluation and ray-centred perturbation theory may wish to simply skim the analytical and algebraic arguments, focusing upon the final asymptotic $\omega \to \infty$ and *ab initio* ray-theoretical results, eqs (53) and (163). These results are compared and discussed in Section 7.

The earliest seismological comparison between the Born approximation and geometrical ray perturbation theory of which we are aware is that conducted for surface waves on a laterally homogeneous background earth model by Snieder (1987). The most relevant previous analyses of body-wave propagation in a 3D earth model are those conducted by Farra & Madariaga (1987), Nowack & Lutter (1988), Coates & Chapman (1990) and Snieder & Lomax (1996).

2 3D KERNELS

Our procedure for determining the 3D Fréchet sensitivity kernel for a body-wave amplitude measurement is analogous—in fact, very nearly identical—to the procedure used to find the 3D traveltime Fréchet kernel by Dahlen *et al.* (2000) and Hung *et al.* (2000), hereinafter referred to as banana–doughnut I and II, respectively. We shall, in this paper, conduct a side-by-side comparison of the traveltime and amplitude kernels in order to highlight the strong similarities as well as the equally stark differences. The notation established in banana–doughnut I will be generally adhered to, although we could not resist a few minor embellishments, which we shall point out as they are introduced. The first such notational difference is that we will conduct the present analysis in terms of the slowness $\sigma = c^{-1}$ rather than the wave speed c. The first-order fractional slowness and wave speed perturbations are, of course, related by $\delta\sigma/\sigma = -\delta c/c$.

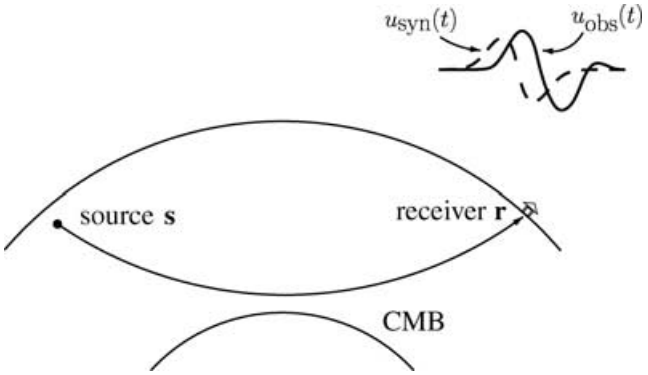

Figure 1. Schematic depiction of the geometrical raypath for a turning P or S wave between a point source **s** (dot) and a receiver **r** (doghouse). In general, the traveltime $T_{\rm obs}$ and the amplitude $A_{\rm obs}$ of an observed pulse $u_{\rm obs}(t)$ (solid line) will differ from the theoretical traveltime $T_{\rm syn}$ and amplitude $A_{\rm syn}$ of the corresponding synthetic pulse $u_{\rm syn}(t)$ (dashed line). The 3D Fréchet kernels K_T and K_A express the sensitivity of the traveltime and logarithmic amplitude perturbations $\delta T = T_{\rm obs} - T_{\rm syn}$ and $\delta(\ln A) = (A_{\rm obs} - A_{\rm syn})/A_{\rm syn}$ to laterally heterogeneous seismic slowness variations $\delta\sigma/\sigma$ within the earth. The theory is applicable to any isolated pulse that propagates between **s** and **r**, including surface reflections or reflections off of the core–mantle boundary.

2.1 The problem

We begin with a general—and admittedly rather vague—statement of the problem. Attention is focused upon a single body-wave phase such as P, PcP, PP, ... or S, ScS, SS, ..., which propagates between a source point **s** and a receiver **r**, as illustrated in Fig. 1. We presume that we have the capability of computing a synthetic waveform $u_{\rm syn}(t)$ for this phase in a background earth model with a spherically symmetric slowness distribution σ. The observed pulse $u_{\rm obs}(t)$ differs from the synthetic $u_{\rm syn}(t)$ because it propagates through the real Earth, which has a slowness distribution $\sigma + \delta\sigma$, where $\delta\sigma$ is a 3D perturbation that we seek to determine. We denote the perturbation in the waveform by

$$\delta u(t) = u_{\rm obs}(t) - u_{\rm syn}(t). \tag{1}$$

The corresponding perturbations in the traveltime and the logarithm of the amplitude of the pulse are defined in an analogous manner, namely,

$$\delta T = T_{\rm obs} - T_{\rm syn}, \tag{2}$$

$$\delta(\ln A) = (A_{\rm obs} - A_{\rm syn})/A_{\rm syn}. \tag{3}$$

Our objective is to find the first-order dependence of the measured traveltime and logarithmic amplitude residuals δT and $\delta(\ln A)$ upon the 3D fractional slowness perturbation $\delta\sigma/\sigma$. Allowing for the possibility that the measured residuals δT and $\delta(\ln A)$ may, in principle, depend upon $\delta\sigma/\sigma$ at every point **x** throughout the 3D earth, we write the desired relations in the general linear form

$$\delta T = \iiint_{\oplus} K_T(\delta\sigma/\sigma)\, d^3\mathbf{x}, \tag{4}$$

$$\delta(\ln A) = \iiint_{\oplus} K_A(\delta\sigma/\sigma)\, d^3\mathbf{x}. \tag{5}$$

Topographic undulations of the outer free surface and any internal discontinuities are not considered, so that the integrals in eqs (4) and (5) are over the spherically symmetric volume $\oplus = \{0 \le \|\mathbf{x}\| \le 6371 \text{ km}\}$. The quantities K_T and K_A are 3D Fréchet kernels expressing the sensitivity of δT and $\delta(\ln A)$ to the volumetric slowness perturbations $\delta\sigma/\sigma$. Subscripts T and A have been introduced to distinguish the traveltime and amplitude kernels K_T and K_A, respectively.

A two-step procedure is required to find the 3D Fréchet sensitivity kernels in eqs (4) and (5):

(1) we must first express each of the seismic observables δT and $\delta(\ln A)$ as a linear functional of the waveform perturbation $\delta u(t)$; and
(2) then we must use the Born approximation to obtain a linear relation between $\delta u(t)$ and the fractional slowness perturbation $\delta\sigma/\sigma$.

We briefly review the outcome of each of these two steps in the case of the traveltime kernel K_T before considering the elastic amplitude kernel K_A, which is the central topic of this paper.

2.2 Traveltime kernel

The traveltime residual δT is presumed, as in banana–doughnut I, to be measured by cross-correlation of the observed and the synthetic pulses:

$$\int_{t_1}^{t_2} u_{\rm syn}(t - \delta T) u_{\rm obs}(t)\, dt = \text{maximum}. \tag{6}$$

Evidently, δT is the amount by which the synthetic pulse $u_{\rm syn}(t)$ must be shifted in time in order to most closely resemble the observed pulse $u_{\rm obs}(t)$ over the time interval $t_1 \le t \le t_2$. A negative traveltime residual, $\delta T < 0$, corresponds to an advance in the arrival of the observed pulse, whereas a positive residual, $\delta T > 0$, corresponds to a delay. Linearization of the cross-correlation measurement criterion (6) leads to an explicit representation of the traveltime shift δT in terms of the waveform perturbation $\delta u(t)$:

$$\delta T = \frac{\int_{t_1}^{t_2} \dot{u}(t)\,\delta u(t)\,dt}{\int_{t_1}^{t_2} \ddot{u}(t) u(t)\,dt}, \tag{7}$$

where a dot denotes differentiation with respect to time, and where we have dropped the identifying subscript upon the unperturbed waveform $u(t) = u_{\rm syn}(t)$ for simplicity. Parseval's relation can be used to express the result (7) in the frequency rather than the time domain:

$$\delta T = \frac{\mathrm{Re} \int_0^\infty i\omega\, u^*(\omega)\,\delta u(\omega)\,d\omega}{\int_0^\infty \omega^2 |u(\omega)|^2\,d\omega}, \tag{8}$$

where an asterisk denotes complex conjugation. The Fourier sign convention in eq. (8) is the same as that in banana–doughnut I: the factor $\exp(-i\omega t)$ appears in the integral in transforming from time t to angular frequency ω, so that $\partial_t \leftrightarrow i\omega$.

Because we have already linearized eq. (6), the Born approximation is the obvious method of choice to find the first-order waveform perturbations $\delta u(t)$ and $\delta u(\omega)$ in eqs (7) and (8). To obtain a convenient representation of the 3D traveltime Fréchet kernel K_T in eq. (4), three additional approximations are needed. These are described in detail and justified in banana–doughnut I; we present a highly abbreviated summary here:

(1) geometrical ray theory is used is used to compute the unperturbed response of the background earth model;
(2) a forward scattering approximation is made at every off-ray scatterer **x**; and
(3) paraxial ray theory is used to describe the response in the vicinity of the central source-to-receiver ray.

We express the position of a scatterer **x** where we wish to compute the sensitivity K_T in the form

$$\mathbf{x} = \boldsymbol{\xi} + \mathbf{q}, \tag{9}$$

where $\boldsymbol{\xi}$ is the nearest point on the central ray, as illustrated in Fig. 2. Points $\boldsymbol{\xi}$ along the ray are parametrized in terms of the arclength $0 \le l \le L$, where the endpoints $l = 0$ and $l = L$ correspond to the source **s** and the receiver **r**, respectively. Upon decomposing the perpendicular offset vector into two orthogonal components,

$$\mathbf{q} = q_1 \hat{\mathbf{e}}_1 + q_2 \hat{\mathbf{e}}_2, \tag{10}$$

where $\hat{\mathbf{e}}_\alpha \cdot \hat{\mathbf{e}}_\beta = \delta_{\alpha\beta}$, we can represent the locations of every scatterer **x** and its associated projection point $\boldsymbol{\xi}$ in terms of their ray-centred coordinates:

$$\mathbf{x} = (q_1, q_2, l), \qquad \boldsymbol{\xi} = (0, 0, l). \tag{11}$$

A key role in the theory is played by the traveltime difference $\Delta T = T_{\rm xs} + T_{\rm xr} - T_{\rm rs}$ required to take the detour path from the source **s** through the scatterer **x** to the receiver **r** rather than travelling directly along the central ray from **s** to **r**. In the paraxial approximation, this difference is a quadratic function of the perpendicular offset vector **q**, of the form

$$\Delta T = \frac{1}{2} \mathbf{q} \cdot \overleftrightarrow{\mathbf{M}} \cdot \mathbf{q}, \quad \text{where} \quad \overleftrightarrow{\mathbf{M}} = \overleftrightarrow{\mathbf{M}}^{\rm T}. \tag{12}$$

The symmetric two-tensor $\overleftrightarrow{\mathbf{M}}$ is the sum of the traveltime Hessians measured in the forward direction from the source **s** and in the backward direction from the receiver **r**:

$$\overleftrightarrow{\mathbf{M}} = \vec{\mathbf{M}} + \overleftarrow{\mathbf{M}}. \tag{13}$$

We have introduced two refinements intended to clarify the notation introduced in banana–doughnut I. First, a sans serif font has been used to express the two-vector eq. (10) and the various traveltime Hessians in eq. (13), with 2×2 component matrices of the form

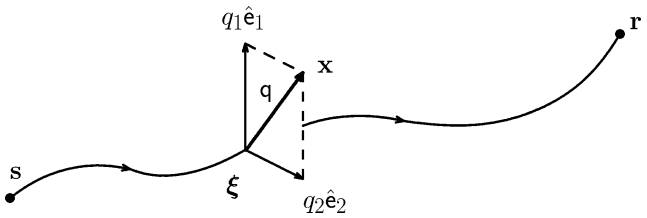

Figure 2. Schematic depiction of a curvaceous geometrical raypath between a source **s** and a receiver **r**. Every off-path scatterer **x** is perpendicularly projected on to the nearest point $\boldsymbol{\xi}$ on the ray. The difference vector is expressed in terms of the two orthogonal shear wave polarizations: $\mathbf{q} = q_1 \hat{\mathbf{e}}_1 + q_2 \hat{\mathbf{e}}_2$. The ray-centred coordinates of the scatterer and its perpendicular projection point are $\mathbf{x} = (q_1, q_2, l)$ and $\boldsymbol{\xi} = (0, 0, l)$, where $0 \le l \le L$ is the arclength along the geometrical ray.

$$\mathbf{M} = \begin{pmatrix} M_{11} & M_{12} \\ M_{12} & M_{22} \end{pmatrix}, \quad \text{where} \quad M_{\alpha\beta} = \frac{\partial^2 T}{\partial q_\alpha \partial q_\beta}. \tag{14}$$

Secondly, forward and backward arrows have been introduced over $\vec{\mathbf{M}}$ and $\overleftarrow{\mathbf{M}}$ to serve as a reminder that they represent the forward and backward Hessians from **s** and **r**, respectively; likewise a double arrow is used over $\overleftrightarrow{\mathbf{M}}$ to indicate that it is the sum of both the forward and backward Hessians. We shall adhere to both the sans serif convention for all inherently 2D ray-centred quantities and the use of forward and backward arrows to indicate the direction that information is propagated along the central ray throughout the remainder of this paper.

The ray-theoretical, forward-scattering, paraxial traveltime Fréchet kernel K_T obtained in banana–doughnut I is given in terms of the forward-plus-backward Hessian $\overleftrightarrow{\mathbf{M}}$ by

$$K_T = \frac{\sigma}{2\pi}\sqrt{|\det \overleftrightarrow{\mathbf{M}}|}\,\frac{\int_0^\infty \omega^3 |u(\omega)|^2 \sin\Phi\, d\omega}{\int_0^\infty \omega^2 |u(\omega)|^2\, d\omega}, \tag{15}$$

where

$$\Phi = \frac{1}{2}\omega\, \mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q} - (\operatorname{sig}\overleftrightarrow{\mathbf{M}} - 2)\frac{\pi}{4}. \tag{16}$$

The quantity $\det\overleftrightarrow{\mathbf{M}}$ is the determinant, whereas $\operatorname{sig}\overleftrightarrow{\mathbf{M}}$ is the signature, or number of positive eigenvalues minus the number of negative eigenvalues, of the two-way Hessian $\overleftrightarrow{\mathbf{M}}$. There are only three possible values of the signature, namely $\operatorname{sig}\overleftrightarrow{\mathbf{M}} = 2, 0, -2$. It is immaterial in the present approximation whether the initial factor of σ in eq. (15) is evaluated at the scatterer **x** or at its perpendicular projection point $\boldsymbol{\xi}$ on the central geometrical ray; we adopt the former alternative for consistency with banana–doughnut I.

2.3 Comment on the generality

In present-day practical applications of 3D Fréchet kernel theory to global seismology, the background slowness distribution σ will be spherically symmetric, as already noted. Geometrical ray paths in that case lie within the plane passing through the source **s**, the receiver **r** and the centre of the Earth, and the associated traveltime Hessians $\vec{\mathbf{M}}$, $\overleftarrow{\mathbf{M}}$ and $\overleftrightarrow{\mathbf{M}}$ of a P, SV or SH wave are diagonal (banana–doughnut I). Eqs (15) and (16) do not, however, depend upon this restriction; they are valid in the ray-theoretical, forward-scattering, paraxial approximation for an arbitrary, piecewise continuous 3D background earth model, with traveltime Hessians $\vec{\mathbf{M}}$, $\overleftarrow{\mathbf{M}}$ and $\overleftrightarrow{\mathbf{M}}$ of the general symmetric form (14). In the interest of maximum generality, we shall continue to adopt this 'not necessarily spherical' point of view throughout the remainder of this paper. The schematic unperturbed ray path depicted in Fig. 2 is obviously of this more general form.

2.4 Amplitude kernel

We define the synthetic and observed body-wave amplitudes to be the rms averages of the corresponding time-domain pulses $u_{\rm syn}(t)$ and $u_{\rm obs}(t)$ over the arrival interval $t_1 \le t \le t_2$:

$$A_{\rm syn} = \sqrt{\frac{1}{t_2 - t_1}\int_{t_1}^{t_2} u_{\rm syn}^2(t)\, dt}, \tag{17}$$

$$A_{\rm obs} = \sqrt{\frac{1}{t_2 - t_1}\int_{t_1}^{t_2} u_{\rm obs}^2(t)\, dt}. \tag{18}$$

Upon inserting the representation $u_{\rm obs}(t) = u_{\rm syn}(t) + \delta u(t)$ into eq. (18) and neglecting second-order terms, we obtain an explicit expression for the logarithmic perturbation in the amplitude $\delta(\ln A) = (A_{\rm obs} - A_{\rm syn})/A_{\rm syn}$, analogous to eq. (7):

$$\delta(\ln A) = \frac{\int_{t_1}^{t_2} u(t)\,\delta u(t)\, dt}{\int_{t_1}^{t_2} u^2(t)\, dt}, \tag{19}$$

where we have again dropped the mnemonic subscript upon the unperturbed pulse $u(t) = u_{\rm syn}(t)$. The corresponding result in the frequency domain analogous to eq. (8) is

$$\delta(\ln A) = \frac{\operatorname{Re}\int_0^\infty u^*(\omega)\,\delta u(\omega)\, d\omega}{\int_0^\infty |u(\omega)|^2\, d\omega}. \tag{20}$$

The amplitude result (20) differs from the corresponding traveltime result (8) only by virtue of the 'missing' factors of $i\omega$ and ω^2 in the numerator and denominator. Upon introducing the ray-theoretical, forward-scattering, paraxial Born approximation to $\delta u(\omega)$, we find that these small differences lead to similarly small differences between the amplitude and traveltime Fréchet kernels. Rather than regurgitating the derivation in banana–doughnut I, we simply state the final result: the 3D amplitude kernel in eq. (5) is given by

$$K_A = \frac{\sigma}{2\pi}\sqrt{|\det \overleftrightarrow{\mathbf{M}}|}\,\frac{\int_0^\infty \omega^2 |u(\omega)|^2 \cos\Phi\, d\omega}{\int_0^\infty |u(\omega)|^2\, d\omega}, \tag{21}$$

where the argument Φ is defined by eq. (16) as before. It is easy to see where the differences between the representations eqs (15) and (21) come from; in particular, the substitution $\sin\Phi \rightarrow \cos\Phi$ arises from the 'missing' factor of i in the numerator of eq. (20). The seemingly minor differences between the 3D Fréchet kernels K_T and K_A lead to major differences in the asymptotic high-frequency behaviour of δT and $\delta(\ln A)$, as we show next.

2.5 Differential kernels

In banana–doughnut I, we generalized the 3D traveltime Fréchet kernel formulation eq. (4) to the case of a differential traveltime, measured by cross-correlation of two observed phases, such as *PcP* and *P* or *ScS* and *S,* at the same station. Distinguishing the two phases by the presence or absence of a prime, we express the residual in the differential traveltime $T' - T$, relative to the background unperturbed earth model, in the form

$$\delta(T' - T) = \iiint_{\oplus} K_{T'-T}(\delta\sigma/\sigma)\, d^3\mathbf{x}. \tag{22}$$

The differential Fréchet kernel $K_{T-T'}$ is then simply the difference of the individual Fréchet kernels, provided that the two pulses $u'(t)$ and $u(t)$ have experienced the same number of caustic passages:

$$K_{T'-T} = K_{T'} - K_T \quad \text{if} \quad M' = M, \tag{23}$$

where M' and M are the associated Maslov indices. In measuring a differential *PP–P* or *SS–S* traveltime, it is common practice to Hilbert transform one or the other phase prior to cross-correlation, in order to eliminate the Maslov inequality $M' \neq M$ (e.g. Kuo *et al.* 1987; Sheehan & Solomon 1991; Woodward & Masters 1991).

A result analogous to eqs (22) and (23) applies to differential amplitude measurements. In this case, the quantity of interest is the ratio A'/A of the rms amplitudes of the primed and unprimed pulses. The residual of the logarithm of this ratio, relative to the background unperturbed earth model, is expressed as

$$\delta[\ln(A'/A)] = \iiint_{\oplus} K_{A'/A}(\delta\sigma/\sigma)\, d^3\mathbf{x}. \tag{24}$$

Since $\ln(A'/A) = \ln A' - \ln A$, it is clear that the differential amplitude Fréchet kernel in eq. (24) is the difference of the individual amplitude kernels, as in eq. (23):

$$K_{A'/A} = K_{A'} - K_A \quad \text{if} \quad M' = M. \tag{25}$$

In principle, the Maslov equality constraint $M' = M$ is not required in this case, since a pulse and its Hilbert transform have the same rms amplitude. Practical *PP*/*P* or *SS*/*S* amplitude measurement schemes are, however, likely to incorporate a Hilbert transformation, since it is preferable to compare the attributes of similarly shaped pulses.

3 HIGH-FREQUENCY LIMIT

In the asymptotic limit of infinite frequency, $\omega \rightarrow \infty$, we expect both of the 3D Born integrals eqs (4) and (5) to reduce to—or at least be consistent with—the corresponding 1D results obtained using ray perturbation theory. Indeed, this consistency could be considered to be a requirement of any properly formulated finite-frequency Fréchet sensitivity kernel. The asymptotic reduction of the 3D traveltime integral eq. (4) to linearized geometrical ray theory was confirmed in banana–doughnut I. We review the relatively simple argument leading to this result here, partly in order to establish some notation, before discussing the more involved high-frequency asymptotic evaluation of the 3D amplitude integral eq. (5). In reducing both eqs (4) and (5) we make use of the Gaussian integral identity

$$\iint_{-\infty}^{\infty} \mathrm{e}^{\frac{1}{2}i\omega\,\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q}}\, d^2\mathbf{q} = \frac{2\pi\,\mathrm{e}^{i(\mathrm{sig}\,\overleftrightarrow{\mathbf{M}})\pi/4}}{\omega\sqrt{|\det\overleftrightarrow{\mathbf{M}}|}}, \quad \omega > 0. \tag{26}$$

The result eq. (26) is exact for any Hessian two-tensor $\overleftrightarrow{\mathbf{M}} = \overleftrightarrow{\mathbf{M}}^{\mathrm{T}}$ with signature $\mathrm{sig}\,\overleftrightarrow{\mathbf{M}} = 2, 0, -2$.

3.1 Reduction of the traveltime integral

We begin by decomposing the 3D integral eq. (4) over the entire Earth $\oplus$ into a 1D integral along the central geometrical ray and a 2D integral over the two transverse coordinates:

$$\iiint_{\oplus} K_T(\delta\sigma/\sigma)\, d^3\mathbf{x} = \int_0^L dl \iint_{-\infty}^{\infty} h K_T(\delta\sigma/\sigma)\, d^2\mathbf{q}. \tag{27}$$

The factor h is the 3D Jacobian of the transformation from Cartesian coordinates $\mathbf{x} = (x_1, x_2, x_3)$ to ray-centred coordinates $\mathbf{x} = (q_1, q_2, l)$, given by Červený & Hron (1980) and Červený (1985)

$$h(q_1, q_2, l) = 1 - \sigma^{-1}\mathbf{q}\cdot\boldsymbol{\nabla}_\perp\sigma, \tag{28}$$

where both the background slowness σ and its cross-path gradient $\boldsymbol{\nabla}_\perp\sigma = \hat{\mathbf{e}}_1(\partial\sigma/\partial q_1) + \hat{\mathbf{e}}_2(\partial\sigma/\partial q_2)$ are evaluated at the point $\boldsymbol{\xi} = (0, 0, l)$ on the central geometrical ray. The infinite limits on the 2D transverse integral are purely formal; in practice, destructive interference between adjacent frequencies ω and $\omega + d\omega$ renders both kernels K_T and K_A negligible except within the first one or two Fresnel zones about the central geometrical ray. Upon inserting the representation (15) into eq. (27) and interchanging the order of $d^2\mathbf{q}$ and $d\omega$ integration, we can express the traveltime perturbation in the form

$$\delta T = \frac{\int_0^L dl \int_0^\infty \omega^3 |u(\omega)|^2 I_T(l, \omega)\, d\omega}{\int_0^\infty \omega^2 |u(\omega)|^2\, d\omega}, \tag{29}$$

where

$$I_T = \frac{1}{2\pi}\sqrt{|\det\overleftrightarrow{\mathbf{M}}|}\iint_{-\infty}^{\infty} h\,\delta\sigma\,\sin\left[\frac{1}{2}\omega\,\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q} - (\operatorname{sig}\overleftrightarrow{\mathbf{M}} - 2)\frac{\pi}{4}\right] d^2\mathbf{q}. \tag{30}$$

To find the leading term in the asymptotic expansion of the quantity I_T in the limit $\omega \to \infty$, it suffices to replace the product multiplying the sinusoid by its lowest-order Taylor series expansion,

$$h(q_1, q_2, l)\,\delta\sigma(q_1, q_2, l) \approx \delta\sigma(0, 0, l), \tag{31}$$

about the central geometrical ray. With the $\mathbf{q}$-independent factor on the right-hand side of eq. (31) extracted, the remaining 2D integral in eq. (30) can readily be evaluated with the aid of the Gaussian identity eq. (26):

$$\iint_{-\infty}^{\infty} \sin\left[\frac{1}{2}\omega\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q} - (\operatorname{sig}\overleftrightarrow{\mathbf{M}} - 2)\frac{\pi}{4}\right] d^2\mathbf{q} = \operatorname{Im}\iint_{-\infty}^{\infty} \mathrm{e}^{\frac{1}{2}i\omega\,\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q}}\mathrm{e}^{-i(\operatorname{sig}\overleftrightarrow{\mathbf{M}}-2)\pi/4}\, d^2\mathbf{q} = \operatorname{Im}\left[\frac{2\pi\,\mathrm{e}^{i\pi/2}}{\omega\sqrt{|\det\overleftrightarrow{\mathbf{M}}|}}\right] = \frac{2\pi}{\omega\sqrt{|\det\overleftrightarrow{\mathbf{M}}|}}. \tag{32}$$

Upon utilizing the result (32) in eq. (30) we find that

$$I_T \approx \omega^{-1}\delta\sigma, \tag{33}$$

where the neglected terms are of order ω^{-2} and higher, and where the slowness perturbation $\delta\sigma$ is evaluated on the central ray $\boldsymbol{\xi} = (0, 0, l)$. Insertion of the $\omega \to \infty$ asymptotic relation (33) into eq. (29) leads to a cancellation of the spectral integrals $\int_0^\infty \omega^2|u(\omega)|^2\, d\omega$, leaving the simple frequency-independent result

$$\delta T \approx \int_0^L \delta\sigma\, dl. \tag{34}$$

The 1D integral eq. (34) is precisely the first-order traveltime shift δT predicted by geometrical ray perturbation theory. Fermat's principle guarantees that the change in the traveltime owing to the change in the ray path between a fixed source $\mathbf{s}$ and receiver $\mathbf{r}$ is of second order in the 3D heterogeneity $\delta\sigma$, so that the integration in eq. (34) may be carried out along the unperturbed ray.

3.2 Reduction of the amplitude integral

We follow a parallel strategy in seeking to reduce the 3D Fréchet amplitude integral eq. (5). Upon transforming from $\mathbf{x} = (x_1, x_2, x_3)$ to $\mathbf{x} = (q_1, q_2, l)$ coordinates and interchanging the order of integration, we are required to evaluate

$$\delta(\ln A) = \frac{\int_0^L dl \int_0^\infty \omega^2 |u(\omega)|^2 I_A(l, \omega)\, d\omega}{\int_0^\infty |u(\omega)|^2\, d\omega}, \tag{35}$$

where

$$I_A = \frac{1}{2\pi}\sqrt{|\det\overleftrightarrow{\mathbf{M}}|}\iint_{-\infty}^{\infty} h\,\delta\sigma\,\cos\left[\frac{1}{2}\omega\,\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q} - (\operatorname{sig}\overleftrightarrow{\mathbf{M}} - 2)\frac{\pi}{4}\right] d^2\mathbf{q}. \tag{36}$$

The difference $\sin\Phi \to \cos\Phi$ between eqs (30) and (36) seems innocent enough; however, upon repeating the argument in eq. (32), we find that

$$\iint_{-\infty}^{\infty} \cos\left[\frac{1}{2}\omega\,\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q} - (\operatorname{sig}\overleftrightarrow{\mathbf{M}} - 2)\frac{\pi}{4}\right] d^2\mathbf{q} = \operatorname{Re}\iint_{-\infty}^{\infty} \mathrm{e}^{\frac{1}{2}i\omega\,\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q}}\mathrm{e}^{-i(\operatorname{sig}\overleftrightarrow{\mathbf{M}}-2)\pi/4}\, d^2\mathbf{q} = \operatorname{Re}\left[\frac{2\pi\,\mathrm{e}^{i\pi/2}}{\omega\sqrt{|\det\overleftrightarrow{\mathbf{M}}|}}\right] = 0. \tag{37}$$

Since the order ω^{-1} term in the asymptotic expansion of I_A vanishes, we are required to carry out the expansion to order ω^{-2}. The first step in this procedure is to extend the Taylor series expansion in eq. (31) up to second order in the cross-path coordinates:

$$h(q_1, q_2, l)\,\delta\sigma(q_1, q_2, l) \approx \delta\sigma + \mathbf{q}\cdot\left(\boldsymbol{\nabla}_\perp\delta\sigma - \sigma^{-1}\delta\sigma\boldsymbol{\nabla}_\perp\sigma\right) + \frac{1}{2}\mathbf{q}\cdot\left(\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\delta\sigma - 2\sigma^{-1}\boldsymbol{\nabla}_\perp\sigma\boldsymbol{\nabla}_\perp\delta\sigma\right)\cdot\mathbf{q}, \tag{38}$$

where the quantities σ and $\delta\sigma$ and all of the cross-path gradients on the right are evaluated on the central ray $\boldsymbol{\xi} = (0, 0, l)$. The terms in eq. (38) that are linear in the offset vector $\mathbf{q}$ do not contribute an order $\omega^{-3/2}$ term to the 2D integral I_A, since

$$\iint_{-\infty}^{\infty} \mathbf{q} \cdot \Big(\boldsymbol{\nabla}_\perp \delta\sigma - \sigma^{-1} \delta\sigma \boldsymbol{\nabla}_\perp \sigma \Big) \cos \left[\frac{1}{2} \omega \, \mathbf{q} \cdot \overleftrightarrow{\mathbf{M}} \cdot \mathbf{q} - (\operatorname{sig} \overleftrightarrow{\mathbf{M}} - 2) \frac{\pi}{4} \right] d^2\mathbf{q} = 0, \tag{39}$$

by virtue of the odd symmetry of the integrand. It is convenient to consolidate the quadratic terms by defining

$$f(\mathbf{q}) = \frac{1}{2} \mathbf{q} \cdot \Big(\boldsymbol{\nabla}_\perp \boldsymbol{\nabla}_\perp \delta\sigma - 2\sigma^{-1} \boldsymbol{\nabla}_\perp \sigma \boldsymbol{\nabla}_\perp \delta\sigma \Big) \cdot \mathbf{q}, \tag{40}$$

where the dependence upon l is understood. The leading-order ω^{-2} approximation to eq. (36) may be written in terms of the function (40) in the form

$$I_A \approx \frac{1}{2\pi} \sqrt{|\det \overleftrightarrow{\mathbf{M}}|} \iint_{-\infty}^{\infty} f(\mathbf{q}) \cos \left[\frac{1}{2} \omega \, \mathbf{q} \cdot \overleftrightarrow{\mathbf{M}} \cdot \mathbf{q} - (\operatorname{sig} \overleftrightarrow{\mathbf{M}} - 2) \frac{\pi}{4} \right] d^2\mathbf{q}. \tag{41}$$

Strictly speaking, the improper 2D cross-path integral in eq. (41) is undefined. We shall evaluate it subject to the implicit understanding that the quadratic function $f(\mathbf{q})$ has been multiplied by a taper that tends to zero sufficiently strongly in the limit $\|\mathbf{q}\| \to \infty$ to guarantee convergence for all real positive frequencies $0 \le \omega \le \infty$. We are free to introduce such a taper by virtue of the negligibility of the amplitude kernel K_A in eq. (5) beyond the first one or two Fresnel zones about the central geometrical ray. The final asymptotic result that we seek must of course be independent of the detailed shape of the taper. To carry out the evaluation of the tapered integral, we introduce the auxiliary function

$$g(\mathbf{q}) = \frac{1}{\omega} \int_{\frac{1}{2}\omega \mathbf{q} \cdot \overleftrightarrow{\mathbf{M}} \cdot \mathbf{q}}^{\infty} \frac{\sin[\zeta - (\operatorname{sig} \overleftrightarrow{\mathbf{M}} - 2)\pi/4]}{\zeta} \, d\zeta, \tag{42}$$

where ζ is a dummy integration variable. The cosinusoid in eq. (41) can be written in terms of this so-called sine integral in the form

$$\cos \left[\frac{1}{2} \omega \, \mathbf{q} \cdot \overleftrightarrow{\mathbf{M}} \cdot \mathbf{q} - (\operatorname{sig} \overleftrightarrow{\mathbf{M}} - 2) \frac{\pi}{4} \right] = -\frac{1}{2} \overleftrightarrow{\mathbf{M}}^{-1} : \boldsymbol{\nabla}_\perp \boldsymbol{\nabla}_\perp g, \tag{43}$$

where the double dot product of two arbitrary two-tensors $\mathbf{B}$ and $\mathbf{C}$ is defined by $\mathbf{B} : \mathbf{C} = B_{\alpha\beta} C_{\alpha\beta}$. The integral that we must evaluate is therefore

$$\mathcal{I} = \iint_{-\infty}^{\infty} f \left(-\frac{1}{2} \overleftrightarrow{\mathbf{M}}^{-1} : \boldsymbol{\nabla}_\perp \boldsymbol{\nabla}_\perp g \right) d^2\mathbf{q}. \tag{44}$$

The double cross-path gradient operator $-\frac{1}{2} \overleftrightarrow{\mathbf{M}}^{-1} : \boldsymbol{\nabla}_\perp \boldsymbol{\nabla}_\perp$ can be shifted from g to f by means of an iterated integration by parts:

$$\mathcal{I} = \iint_{-\infty}^{\infty} \left(-\frac{1}{2} \overleftrightarrow{\mathbf{M}}^{-1} : \boldsymbol{\nabla}_\perp \boldsymbol{\nabla}_\perp f \right) g \, d^2\mathbf{q}. \tag{45}$$

The taper is presumed to tend to zero rapidly enough to eliminate any contribution from the boundary at $\|\mathbf{q}\| \to \infty$. The new doubly differentiated factor

$$-\frac{1}{2} \overleftrightarrow{\mathbf{M}}^{-1} : \boldsymbol{\nabla}_\perp \boldsymbol{\nabla}_\perp f = -\frac{1}{2} \overleftrightarrow{\mathbf{M}}^{-1} : \Big(\boldsymbol{\nabla}_\perp \boldsymbol{\nabla}_\perp \delta\sigma - 2\sigma^{-1} \boldsymbol{\nabla}_\perp \sigma \boldsymbol{\nabla}_\perp \delta\sigma \Big) \tag{46}$$

is independent of the cross-path vector $\mathbf{q}$, so it can be taken outside of the integral eq. (45), leaving the result

$$\mathcal{I} = -\frac{1}{2} \overleftrightarrow{\mathbf{M}}^{-1} : \Big(\boldsymbol{\nabla}_\perp \boldsymbol{\nabla}_\perp \delta\sigma - 2\sigma^{-1} \boldsymbol{\nabla}_\perp \sigma \boldsymbol{\nabla}_\perp \delta\sigma \Big) \iint_{-\infty}^{\infty} g(\mathbf{q}) \, d^2\mathbf{q}. \tag{47}$$

Now that the boundary terms have been suppressed, the presence of the $\|\mathbf{q}\| \to \infty$ taper can be ignored. The untapered 2D integral in eq. (47) can be written in the form

$$\begin{aligned} \iint_{-\infty}^{\infty} g(\mathbf{q}) \, d^2\mathbf{q} &= \frac{1}{\omega} \iint_{-\infty}^{\infty} d^2\mathbf{q} \int_{\frac{1}{2}\omega \mathbf{q} \cdot \overleftrightarrow{\mathbf{M}} \cdot \mathbf{q}}^{\infty} \frac{\sin[\zeta - (\operatorname{sig} \overleftrightarrow{\mathbf{M}} - 2)\pi/4]}{\zeta} \, d\zeta = \frac{1}{\omega} \operatorname{Im} \left[\mathrm{e}^{-i(\operatorname{sig} \overleftrightarrow{\mathbf{M}} - 2)\pi/4} \iint_{-\infty}^{\infty} d^2\mathbf{q} \int_{\frac{1}{2}\omega \mathbf{q} \cdot \overleftrightarrow{\mathbf{M}} \cdot \mathbf{q}}^{\infty} \frac{\mathrm{e}^{i\zeta}}{\zeta} \, d\zeta \right] \\ &= \frac{1}{\omega} \operatorname{Re} \left[\mathrm{e}^{-i(\operatorname{sig} \overleftrightarrow{\mathbf{M}})\pi/4} \iint_{-\infty}^{\infty} d^2\mathbf{q} \int_{\frac{1}{2}\omega \mathbf{q} \cdot \overleftrightarrow{\mathbf{M}} \cdot \mathbf{q}}^{\infty} \frac{\mathrm{e}^{i\zeta}}{\zeta} \, d\zeta \right]. \end{aligned} \tag{48}$$

Upon substituting $\zeta = (\frac{1}{2}\omega \, \mathbf{q} \cdot \overleftrightarrow{\mathbf{M}} \cdot \mathbf{q})\eta$ and interchanging the order of $d^2\mathbf{q}$ and $d\eta$ integration, we obtain

$$\iint_{-\infty}^{\infty} d^2\mathbf{q} \int_{\frac{1}{2}\omega \mathbf{q} \cdot \overleftrightarrow{\mathbf{M}} \cdot \mathbf{q}}^{\infty} \frac{\mathrm{e}^{i\zeta}}{\zeta} \, d\zeta = \int_1^{\infty} \frac{d\eta}{\eta} \iint_{-\infty}^{\infty} \mathrm{e}^{\frac{1}{2} i\omega\eta \, \mathbf{q} \cdot \overleftrightarrow{\mathbf{M}} \cdot \mathbf{q}} \, d^2\mathbf{q} = \frac{2\pi \, \mathrm{e}^{i(\operatorname{sig} \overleftrightarrow{\mathbf{M}})\pi/4}}{\omega \sqrt{|\det \overleftrightarrow{\mathbf{M}}|}} \int_1^{\infty} \frac{d\eta}{\eta^2} = \frac{2\pi \, \mathrm{e}^{i(\operatorname{sig} \overleftrightarrow{\mathbf{M}})\pi/4}}{\omega \sqrt{|\det \overleftrightarrow{\mathbf{M}}|}}. \tag{49}$$

Putting all of the above results together, we find that

$$\iint_{-\infty}^{\infty} g(\mathbf{q})\, d^2\mathbf{q} = \frac{2\pi}{\omega^2\sqrt{|\det \overleftrightarrow{\mathbf{M}}|}}, \tag{50}$$

so that

$$\mathcal{I} = -\frac{\pi \overleftrightarrow{\mathbf{M}}^{-1} : \left(\boldsymbol{\nabla}_\perp \boldsymbol{\nabla}_\perp \delta\sigma - 2\sigma^{-1}\boldsymbol{\nabla}_\perp\sigma\boldsymbol{\nabla}_\perp\delta\sigma\right)}{\omega^2\sqrt{|\det \overleftrightarrow{\mathbf{M}}|}} \tag{51}$$

and, therefore,

$$I_A \approx -\frac{1}{2}\omega^{-2}\overleftrightarrow{\mathbf{M}}^{-1} : \left(\boldsymbol{\nabla}_\perp \boldsymbol{\nabla}_\perp \delta\sigma - 2\sigma^{-1}\boldsymbol{\nabla}_\perp\sigma\boldsymbol{\nabla}_\perp\delta\sigma\right). \tag{52}$$

Upon inserting the result eq. (52) into eq. (35) we see that the factor of ω^{-2} is precisely what is needed to give rise to a cancellation of the spectral integrals $\int_0^\infty |u(\omega)|^2\, d\omega$ in the numerator and denominator. The final frequency-independent expression for the logarithmic amplitude perturbation, analogous to the Fermat result eq. (34) for the traveltime, is

$$\delta(\ln A) \approx -\frac{1}{2}\int_0^L \overleftrightarrow{\mathbf{M}}^{-1} : \left(\boldsymbol{\nabla}_\perp \boldsymbol{\nabla}_\perp \delta\sigma - 2\sigma^{-1}\boldsymbol{\nabla}_\perp\sigma\boldsymbol{\nabla}_\perp\delta\sigma\right) dl. \tag{53}$$

Both of the $\omega \to \infty$ asymptotic approximations eqs (34) and (53) are 1D line integrals along the central geometrical ray; however, $\delta(\ln A)$ depends not simply upon the local slowness perturbation $\delta\sigma$, but rather upon its cross-path gradient $\boldsymbol{\nabla}_\perp\delta\sigma$ and curvature $\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\delta\sigma$. In the case of greatest interest, whenever the background slowness σ is much smoother than the 3D perturbation $\delta\sigma$, the dominant contribution will come from the most highly differentiated term, $\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\delta\sigma$. Coates & Chapman (1990) carried out an asymptotic evaluation similar to that in this section, incorrectly omitting the scalefactor h in eq. (36), and restricting attention to the case of a diagonal traveltime Hessian $\overleftrightarrow{\mathbf{M}}$; the above analysis generalizes their result to the case of a general background medium, in which $\overleftrightarrow{\mathbf{M}}$ is of the form eq. (14).

4 ROGUES' GALLERY OF KERNELS

Before proceeding to compare the above asymptotic results with the 1D results obtained using *ab initio* geometrical ray perturbation theory, we pause for a brief pictorial interlude, in which we illustrate the geometry of a number of 3D Fréchet traveltime and amplitude kernels K_T and K_A. The synthetic pulse shape $u_{\rm syn}(t)$ for all of the kernels displayed here is an attenuated ($T^\star = 4$ s) Dirac delta function, convolved with the Albuquerque, New Mexico Seismic Research Observatory (ANMO SRO) instrument response, as used in the shear wave cross-correlation traveltime measurement programme conducted by investigators at the Scripps Institute of Oceanography (Masters *et al.* 1996). The characteristic period of the cross-correlated *S, ScS* and *SS* waves is approximately 20–25 s, as shown in Fig. 3. The two-way traveltime Hessian $\overleftrightarrow{\mathbf{M}}$ in eqs (15) and (21) has been computed using the spherical-earth procedure outlined in banana–doughnut I; the background model adopted for this purpose is PREM (Dziewonski & Anderson 1981). The source $\mathbf{s}$ and the receiver $\mathbf{r}$ in all of our examples are situated on the surface of the Earth, at a radius = 6371 km.

Fig. 4 shows a number of 2D cross-sections through the 3D 'banana–doughnut' traveltime kernel K_T for a turning S wave at an angular epicentral distance $\Delta = 60^\circ$. The presence of the sinusoidal term $\sin\Phi = \sin\frac{1}{2}\omega\,\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q}$ in eq. (15) renders the traveltime sensitivity identically zero everywhere along the central geometrical ray. The maximum sensitivity is in the red 'banana skin' or 'doughnut' in the outer part of the first Fresnel zone, where $K_T > 0$. There is a light-green sideband of negative sensitivity, $K_T < 0$, within the second Fresnel zone, beyond which the sensitivity is negligible. Eq. (32) stipulates that the cross-sectional 'area' of the 3D traveltime kernel is

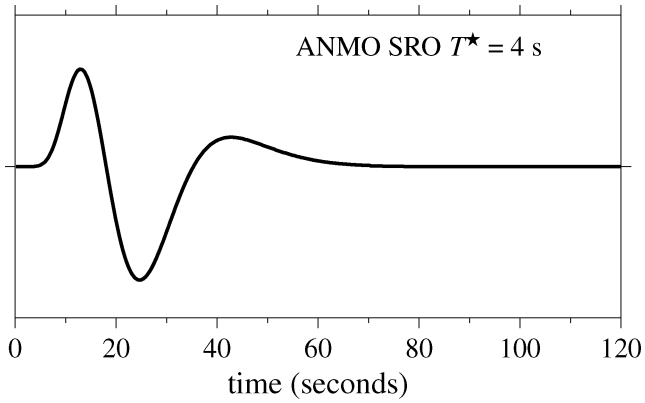

Figure 3. Synthetic pulse shape $u(t) = u_{\rm syn}(t)$ used to compute the 3D Fréchet sensitivity kernels K_T and K_A depicted in Figs 4–9. The input attenuated Dirac delta function is defined by $\delta^\star(t) = \frac{1}{\pi}\mathrm{Re}\int_0^\infty \exp i\omega[t + \frac{1}{2}iT^\star + \frac{1}{\pi}T^\star\ln(\omega/\omega_0)]\, d\omega$, where $\omega_0/2\pi = 1$ Hz is the elastic reference frequency for PREM. The selected attenuation time $T^\star = 4$ s is characteristic of teleseismic *S*, *ScS* and *SS* waves.

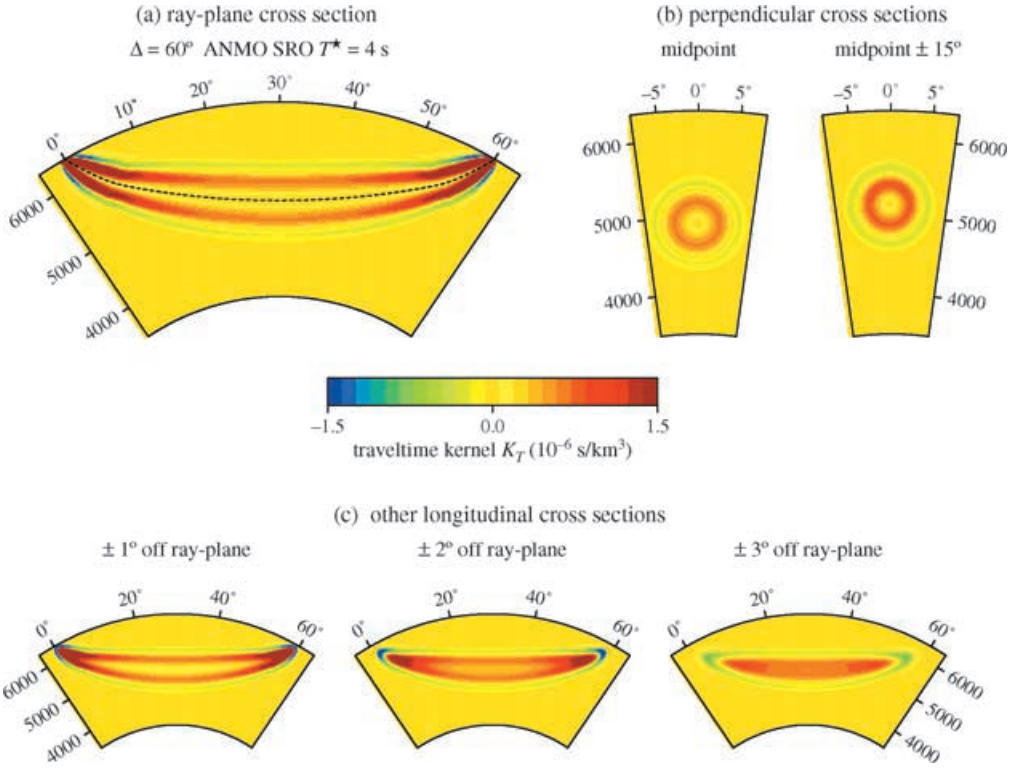

Figure 4. 2D cross-sections through the 3D traveltime Fréchet kernel K_T for a long-period S wave observed at an angular epicentral distance $\Delta = 60^\circ$ in PREM. (a) Ray-plane cross-section; dashed line indicates location of geometrical raypath. (b) Vertical cross-sections perpendicular to the raypath, through the source–receiver midpoint (left) and $\pm 15^\circ$ on either side (right). (c) Vertical cross-sections parallel to the ray plane, at distances $\pm 1^\circ$ (left), $\pm 2^\circ$ (middle) and $\pm 3^\circ$ (right). Note the colour scale: red denotes a region of positive sensitivity, $K_T > 0$, where a slow anomaly, $\delta\sigma > 0$, increases the traveltime, $\delta T > 0$; blue-green denotes a region of negative sensitivity, $K_T < 0$, where a slow anomaly, $\delta\sigma > 0$, decreases the traveltime, $\delta T < 0$; yellow denotes a region of negligible sensitivity, $K_T \approx 0$.

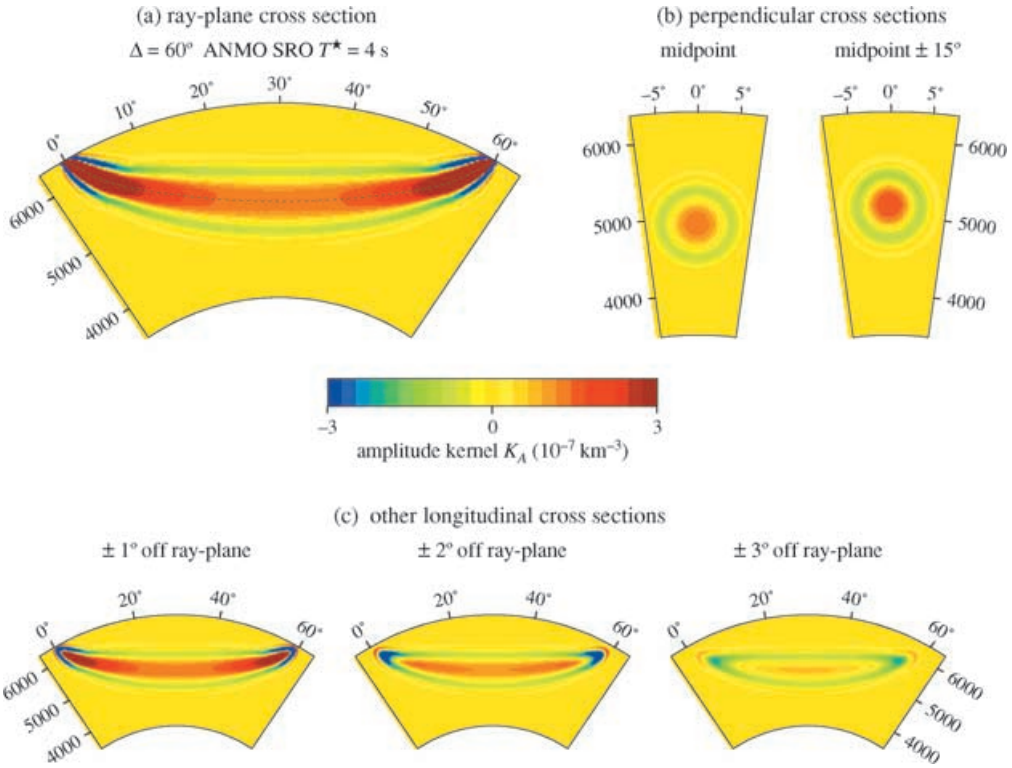

Figure 5. Same as Fig. 4, except for the 3D amplitude Fréchet kernel K_A. Note the difference in units: $|K_T| \leq 1.5 \times 10^{-6}\,\mathrm{s\,km^{-3}}$, whereas $|K_A| \leq 3 \times 10^{-7}\,\mathrm{km^{-3}}$.

$$\int\!\!\int_{-\infty}^{\infty} K_T \, d^2\mathbf{q} \approx \sigma(0, 0, l), \tag{54}$$

correct to order ω^{-1} at every point $0 \le l \le L$ along the source-to-receiver ray. This is the fundamental reason that a cross-correlation traveltime residual (4) reduces to the familiar 1D Fermat integral eq. (34) in the asymptotic limit of infinite frequency, $\omega \to \infty$.

The 3D elastic amplitude Fréchet kernel K_A for the same $\Delta = 60^\circ$ *S* wave is shown in Fig. 5. The amplitude sensitivity exhibits a local maximum, $K_A > 0$, everywhere along the central geometrical ray, by virtue of the presence of the cosinusoidal term $\cos\Phi = \cos\frac{1}{2}\omega\,\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q}$ in eq. (21). The cross-sectional 'area' of the paraxial, forward-scattering amplitude kernel is seen from eq. (37) to be

$$\int\!\!\int_{-\infty}^{\infty} K_A \, d^2\mathbf{q} \approx 0, \tag{55}$$

correct to the same order ω^{-1} at every point $0 \le l \le L$ along the ray. The contrast between eqs (54) and (55) is notable, and is fundamentally responsible for the more complicated dependence of the $\omega \to \infty$ asymptotic result (53) upon the first and second cross-path gradients of the slowness perturbation, rather than simply upon the perturbation itself, as in eq. (34). The manner in which the equality (55) is achieved is evident in the two perpendicular cross-sections: clearly the red ($K_A > 0$) 'jelly' must be cancelled by the surrounding green ($K_A < 0$) 'doughnut'. Any 3D anomaly $\delta\sigma/\sigma$ that is constant across the width of the kernel K_A gives rise to a negligible amplitude perturbation, $\delta(\ln A) \approx 0$, as a consequence of this cancellation. For this reason, body-wave amplitude measurements are more sensitive to small-scale than to large-scale slowness variations $\delta\sigma$. In the $\omega \to \infty$ asymptotic limit, this increased small-scale sensitivity is reflected in the dependence upon $\boldsymbol{\nabla}_\perp\delta\sigma$ and $\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\delta\sigma$.

Cross-sectional views of the traveltime Fréchet kernel eq. (15) and the amplitude kernel eq. (21) for an *ScS* wave at an epicentral distance of $\Delta = 48^\circ$ are depicted in Figs 6 and 7, respectively. Both the 'hollow' K_T 'banana' and the 'filled' K_A 'banana' are folded at the core–mantle boundary. The angular width of the region of non-negligible sensitivity is 20°–30° within the D″ zone.

Fig. 8 shows the traveltime Fréchet kernel K_T for an *SS* wave at an epicentral distance $\Delta = 120^\circ$. The 'pinchoff' points in the ray-plane cross-section at a radius of ∼5000 km and epicentral distances of ∼40° and ∼80° are caustics for the receiver-to-source and the source-to-receiver waves, respectively. The determinant of the two-way traveltime Hessian diverges, $|\det\overleftrightarrow{\mathbf{M}}| \to \infty$, and the signature is discontinuous at these points:

$$\operatorname{sig}\overleftrightarrow{\mathbf{M}} = \begin{cases} 2 & \text{if} \quad 0^\circ \lesssim \Delta \lesssim 40^\circ \\ 0 & \text{if} \quad 40^\circ \lesssim \Delta \lesssim 80^\circ \\ 2 & \text{if} \quad 80^\circ \lesssim \Delta \lesssim 120^\circ. \end{cases} \tag{56}$$

This gives rise to the characteristic zero-to-maximal-to-zero traveltime sensitivity along the central geometrical ray:

$$\sin\frac{1}{2}\omega\,\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q} \to \cos\frac{1}{2}\omega\,\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q} \to \sin\frac{1}{2}\omega\,\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q}. \tag{57}$$

The distinctive saddle-shaped geometry of the 3D kernel K_T in the vicinity of the surface reflection point is a consequence of the intrinsic minimax character of the *SS* phase. Waves that scatter off of an in-plane slowness heterogeneity to the 'left', 'right' or below the 60° bounce point arrive earlier than the background geometrical pulse, whereas those that take a detour to an out-of-plane scatterer arrive later. The cross-path sections at 40° and 80° and the horizontal slice at depth = 1360 km both pass approximately through the caustics. Note that the kernel K_T is only 'pinched' off within the ray plane; the width of the sensitivity in the perpendicular direction is governed by the purely geometrical spreading of cylindrically symmetrical waves within a sphere (banana–doughnut I).

In Fig. 9 we show the amplitude Fréchet kernel K_A for the same $\Delta = 120^\circ$ *SS* wave. In this case the sensitivity varies from maximal to identically zero to maximal along the central geometrical ray:

$$\cos\frac{1}{2}\omega\,\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q} \to -\sin\frac{1}{2}\omega\,\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q} \to \cos\frac{1}{2}\omega\,\mathbf{q}\cdot\overleftrightarrow{\mathbf{M}}\cdot\mathbf{q}. \tag{58}$$

Note that $K_A > 0$ within the ray plane to the 'left', 'right' and below the 60° surface bounce point, whereas $K_A < 0$ in the vicinity of the bounce point, but out of the ray plane. These regions of positive and negative sensitivity are separated by a yellow 'wishbone' on the perpendicular cross-section through the midpoint and by a yellow 'X' on the near-surface (depth = 40 km) section. The equality (55) is achieved between ∼40° and ∼80° by the cancellation of the red and green regions on the 'uphill' ($K_A > 0$) and 'downhill' ($K_A < 0$) sides of the saddle.

5 GEOMETRICAL RAY THEORY

In the case of a traveltime perturbation δT, it is evident that the $\omega \to \infty$ approximation (34) to the 3D integral (4) coincides with the result obtained using geometrical ray perturbation theory. A number of investigators, including Thomson (1983), Nowack & Lutter (1988), Coates & Chapman (1990), Neele *et al.* (1993) and Liu & Tromp (1996) have presented numerical recipes for computing first-order ray-theoretical amplitude perturbations; however, we are not aware of a satisfactory explicit general formula for $\delta(\ln A)$ with which the 1D asymptotic integral representation (53) can be compared. For this reason, we present our own strictly ray-theoretical amplitude analysis in this and the following section. In the interest of completeness, and in order to establish a consistent notation, we shall re-derive a number of well-known results that

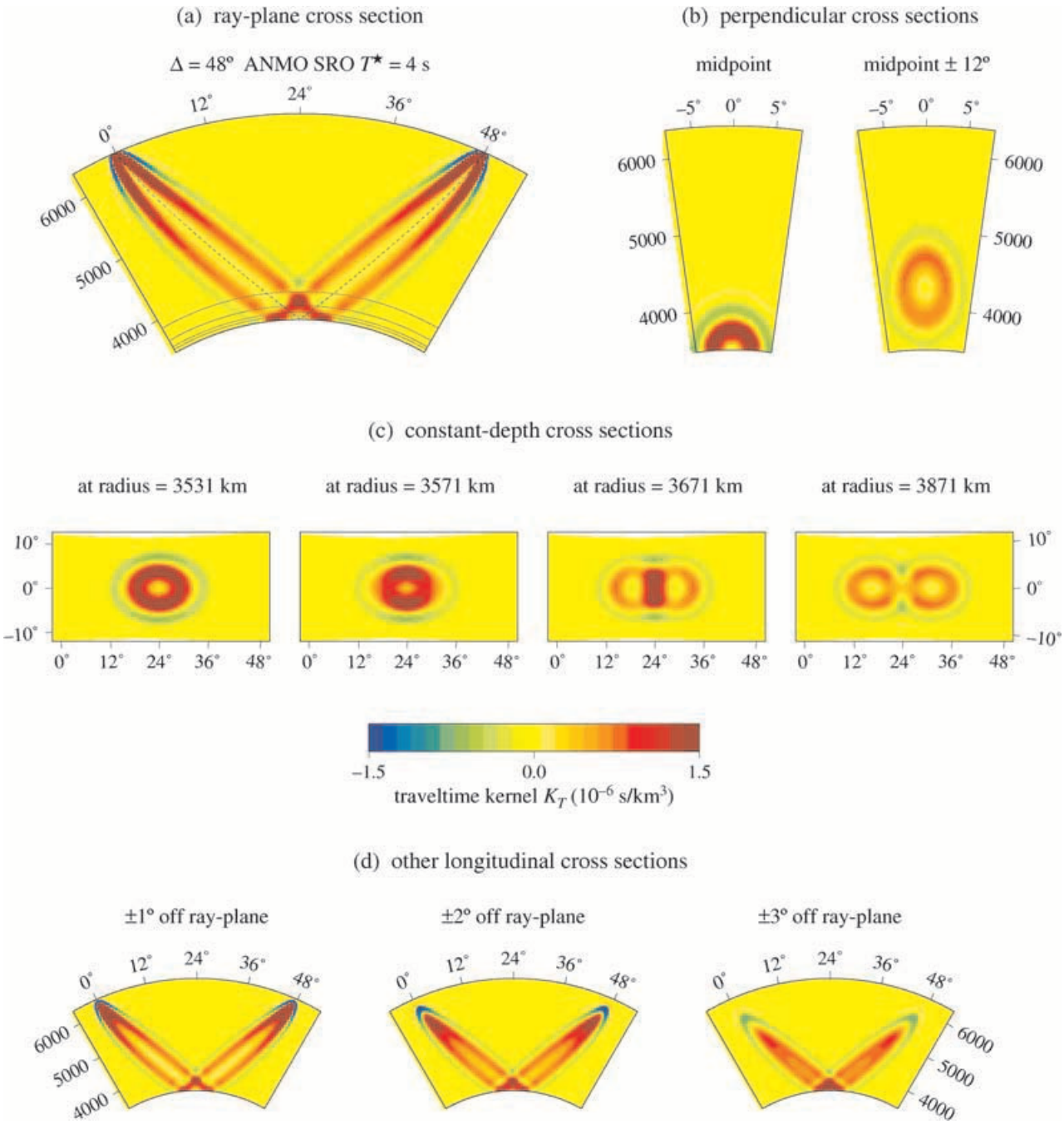

Figure 6. 2D cross-sections through the 3D traveltime Fréchet kernel K_T for an *ScS* core reflection observed at an angular epicentral distance $\Delta = 48^\circ$ in PREM. (a) Ray-plane cross-section; dashed line indicates location of geometrical raypath. Solid lines denote locations of four constant-depth cross-sections. (b) Vertical cross-sections perpendicular to the raypath, through the core–mantle boundary bounce point (left) and $\pm 12^\circ$ on either side (right). (c) Constant-depth cross-sections at radii 3531, 3571, 3671 and 3871 km (left to right). The distances above the core–mantle boundary are 51, 91, 191 and 391 km, respectively; the locations of these slices are indicated on the ray-plane cross-section above. (d) Vertical cross-sections parallel to the ray plane, at distances $\pm 1^\circ$ (left), $\pm 2^\circ$ (middle) and $\pm 3^\circ$ (right).

may be found in standard references, e.g. the authoritative review of seismic ray theory by Červený (1985). Many other ray-theoretical results whose derivations are lengthy or tedious will, however, simply be stated without proof.

5.1 Kinematic ray tracing

We begin by reviewing the problem of tracing rays in the unperturbed background medium. Attention is restricted to the case of an *infinite, smooth* slowness distribution σ, with no internal discontinuities. The position and slowness vectors along a ray will be denoted by $\boldsymbol{\xi}$ and $\boldsymbol{\sigma} = \boldsymbol{\nabla} T$, respectively; the fundamental relation governing ray kinematics is the eikonal equation (Whitham 1974, Section 7.7):

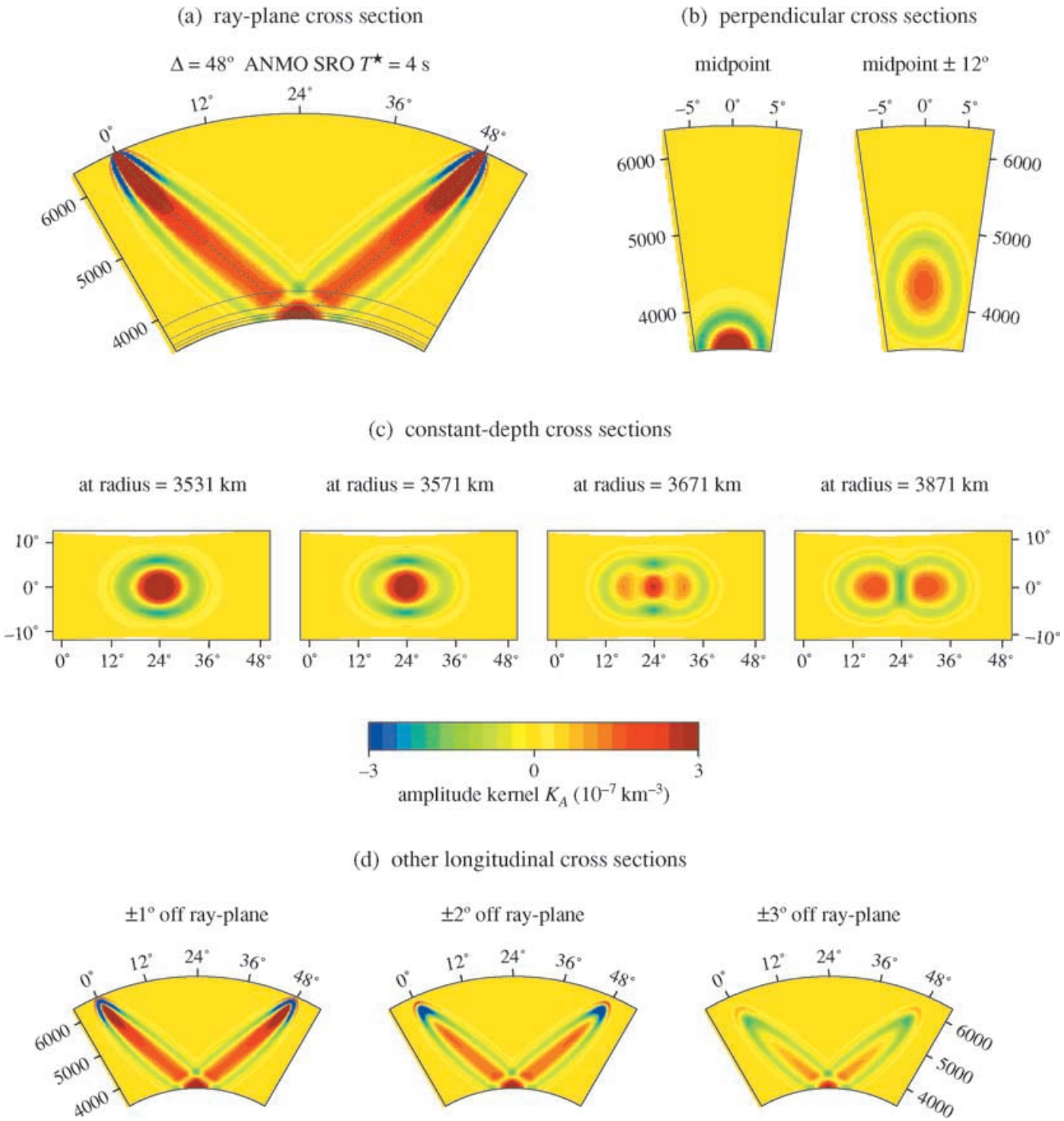

Figure 7. Same as Fig. 6, except for the 3D amplitude Fréchet kernel K_A. Note the difference in units: $|K_T| \leq 1.5 \times 10^{-6}$ s km^{-3} whereas $|K_A| \leq 3 \times 10^{-7}$ km^{-3}.

$$\boldsymbol{\sigma} \cdot \boldsymbol{\sigma} = \boldsymbol{\nabla} T \cdot \boldsymbol{\nabla} T = \sigma^2. \tag{59}$$

Eq. (59) is a first-order, non-linear partial differential equation that can be solved for the traveltime T by the method of characteristics. The variation of position $\boldsymbol{\xi}$ and slowness $\boldsymbol{\sigma}$ along the characteristics or rays is governed by the first-order, linear, ordinary differential equations

$$d\boldsymbol{\xi}/dl = \sigma^{-1}\boldsymbol{\sigma}, \qquad d\boldsymbol{\sigma}/dl = \boldsymbol{\nabla}\sigma. \tag{60}$$

The independent variable l is the arclength along the ray. To trace or 'shoot' a ray from a point source **s**, we must integrate the characteristic eq. (60), subject to the initial conditions

$$\boldsymbol{\xi}(0) = \mathbf{s}, \qquad \boldsymbol{\sigma}(0) = \boldsymbol{\sigma}_s = \sigma_s \hat{\boldsymbol{\sigma}}_s, \tag{61}$$

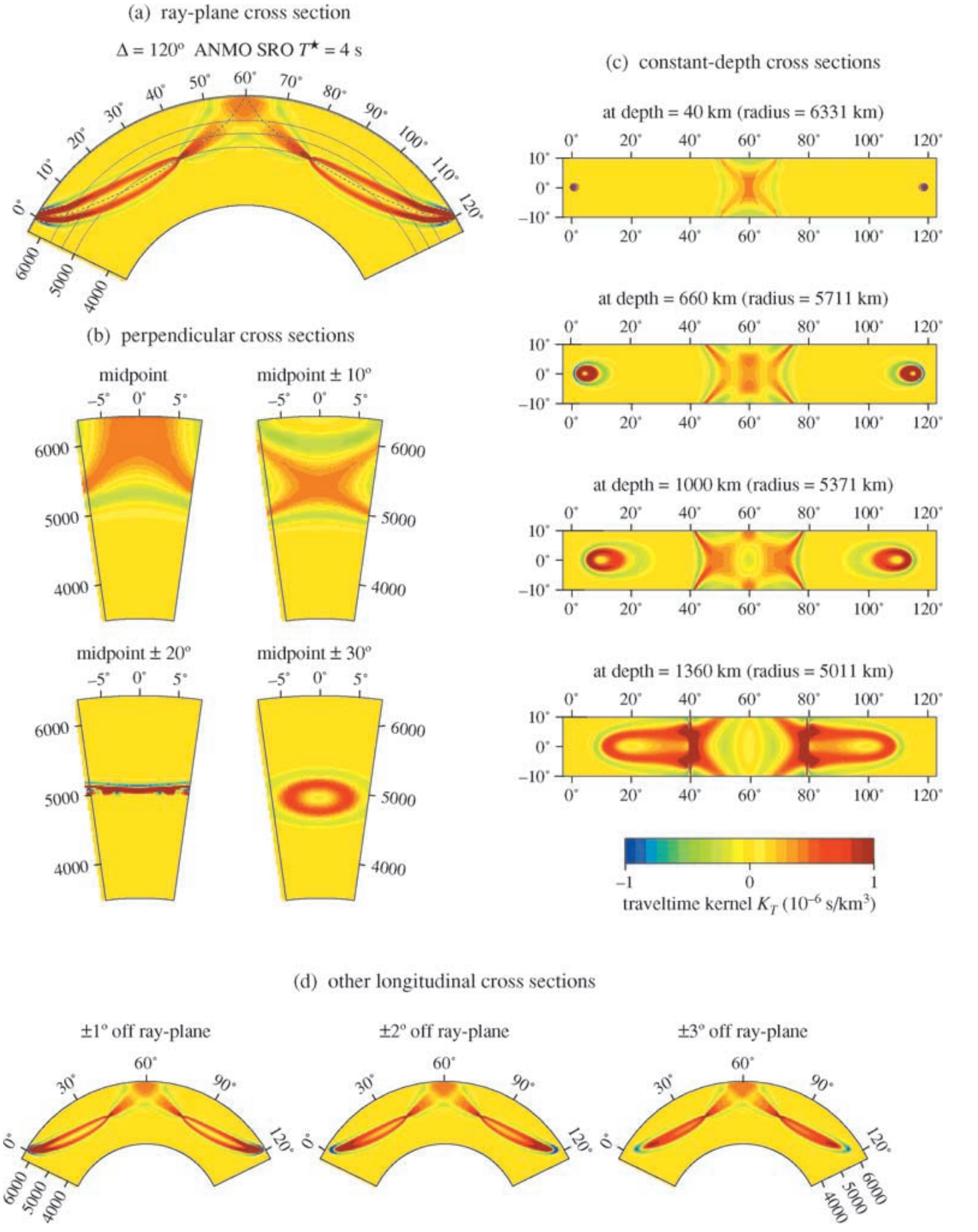

Figure 8. 2D cross-sections through the 3D traveltime Fréchet kernel K_T for an SS surface reflection observed at an angular epicentral distance $\Delta = 120^\circ$ in PREM. (a) Ray-plane cross-section; the dashed line indicates the location of the geometrical raypath. Solid lines denote locations of four constant-depth cross-sections. (b) Vertical cross-sections perpendicular to the raypath, through the surface bounce point (top left) and $\pm 10^\circ$ (top right), $\pm 20^\circ$ (bottom left) and $\pm 30^\circ$ (bottom right) on either side. (c) Constant-depth cross-sections at radii 6331 km (top), 5711 km (second from top), 5371 km (third from top) and 5011 km (bottom). The corresponding depths are 40, 660, 1000 and 1360 km; the locations of these slices are indicated on the ray-plane cross-section. (d) Vertical cross-sections parallel to the ray plane, at distances $\pm 1^\circ$ (left), $\pm 2^\circ$ (middle) and $\pm 3^\circ$ (right).

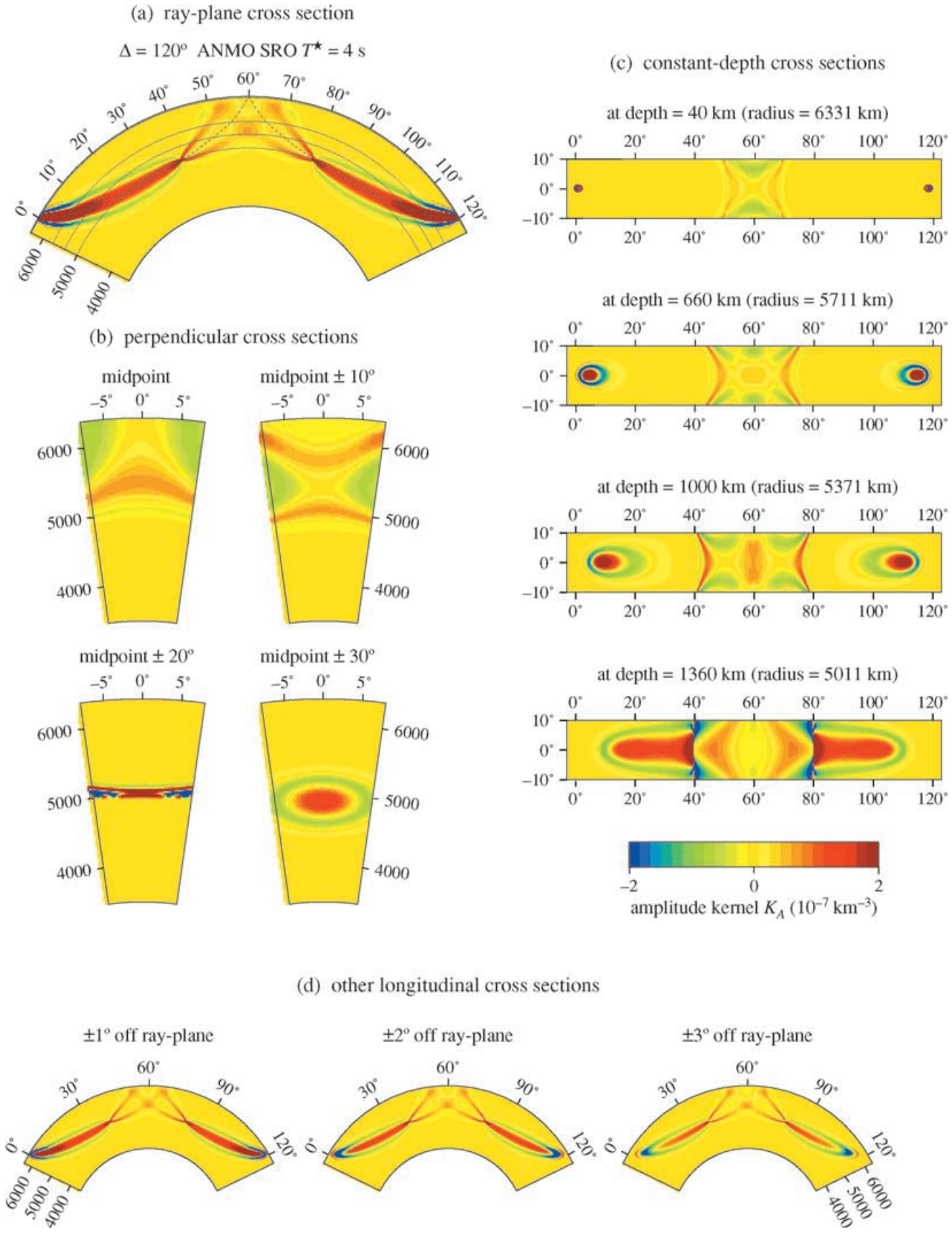

Figure 9. Same as Fig. 8, except for the 3D amplitude Fréchet kernel K_A. Note the difference in units: $|K_T| \leq 1.5 \times 10^{-6}$ s km^{-3}, whereas $|K_A| \leq 3 \times 10^{-7}$ km^{-3}. The ±20° cross-path sections and the 1360 km depth slice both pass approximately through the ~40° receiver-to-source and the ~80° source-to-receiver caustics.

where a circumflex is used to identify a unit vector and a subscript s denotes evaluation at the source. Iteration is required to find the initial takeoff direction $\hat{\boldsymbol{\sigma}}_s$ that enables a ray to 'hit' a receiver $\mathbf{r}$; the endpoint condition that must be satisfied to solve this two-point ray-tracing problem is

$$\boldsymbol{\xi}(L) = \mathbf{r}, \tag{62}$$

where L is the total arclength of the ray between $\mathbf{s}$ and $\mathbf{r}$, as before.

5.2 Cartesian Hamiltonian

The ray tracing eqs (60) are the canonical equations,

$$d\boldsymbol{\xi}/dl = \partial H/\partial \boldsymbol{\sigma}, \qquad d\boldsymbol{\sigma}/dl = -\partial H/\partial \boldsymbol{\xi}, \tag{63}$$

for the conservative Hamiltonian (Whitham 1974; Section 7.7; Farra & Madariaga 1987)

$$H(\boldsymbol{\xi}, \boldsymbol{\sigma}) = \frac{1}{2}\sigma^{-1}(\boldsymbol{\sigma}\cdot\boldsymbol{\sigma} - \sigma^2) = 0. \tag{64}$$

Hamilton's eqs (63) describe the evolution of a ray trajectory in a 6D phase space; the 3D Cartesian position and slowness vectors $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)$ and $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ play the roles of the generalized coordinates and conjugate momenta, respectively. The conservation relation (64) is obviously equivalent to eq. (59) so it can be considered to be the Cartesian Hamiltonian version of the eikonal equation.

5.3 Shear wave polarization

The particle motions associated with a P and an S wave are parallel and perpendicular to the associated ray, respectively. There are two independent shear wave polarizations, $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$, which we take to be defined in such a way that $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3 = \hat{\boldsymbol{\sigma}}$ comprise a right-handed orthonormal triad, satisfying $(\hat{\mathbf{e}}_1 \times \hat{\mathbf{e}}_2)\cdot\hat{\mathbf{e}}_3 = 1$, as illustrated in Fig. 10. In a general 3D medium, the polarization of a shear wave twists around the ray as the wave propagates; the two polarization vectors $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ satisfy independent evolution equations of the form (Červený 1985)

$$d\hat{\mathbf{e}}_\alpha/dl = -\sigma^{-1}(\hat{\mathbf{e}}_\alpha\cdot\boldsymbol{\nabla}_\perp\sigma)\hat{\mathbf{e}}_3, \quad \alpha = 1, 2, \tag{65}$$

where $\boldsymbol{\nabla}_\perp = \boldsymbol{\nabla} - \hat{\mathbf{e}}_3(\hat{\mathbf{e}}_3\cdot\boldsymbol{\nabla})$ denotes the cross-path gradient, as before. To obtain a complete solution to the kinematic ray-tracing problem, it is necessary to solve eqs (65) as well as (60). The initial polarization directions at the source $\mathbf{s}$ may be chosen arbitrarily, subject to the orthonormality constraint $\hat{\mathbf{e}}_\alpha(0)\cdot\hat{\mathbf{e}}_\beta(0) = \delta_{\alpha\beta}$.

In a spherically symmetric earth model, the two shear wave polarizations coincide with the ray normal $\hat{\boldsymbol{\nu}}$ and the binormal $\hat{\boldsymbol{\sigma}}\times\hat{\boldsymbol{\nu}}$, respectively; the SV polarization $\hat{\mathbf{e}}_1 = \hat{\mathbf{e}}_{SV}$ lies in the ray plane and changes direction in a seismologically familiar fashion as the ray turns, whereas the SH polarization $\hat{\mathbf{e}}_2 = \hat{\mathbf{e}}_{SH}$ is perpendicular to the ray plane and constant along the ray. Both SV and SH pulse amplitudes can be measured using a procedure analogous to the one in eqs (17) and (18) and interpreted using eqs (5) and (21) or (53).

5.4 Ray-centred coordinates

We regard the three unit vectors $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3$ as the basis vectors of an orthogonal curvilinear system of ray-centred coordinates. Points $\mathbf{x}$ and $\boldsymbol{\xi}$ off and on of the central ray are represented in this coordinate system by $\mathbf{x} = (q_1, q_2, l)$ and $\boldsymbol{\xi} = (0, 0, l)$, respectively. The scalefactors associated with the three coordinates q_1, q_2, l are (Červený 1985)

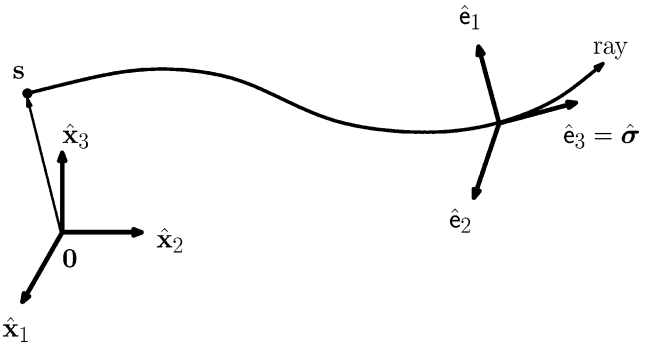

Figure 10. The first-order differential eqs (60) and initial conditions (61) govern the evolution of a geometrical ray that departs from a point source $\mathbf{s}$ in a Cartesian axis system $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \hat{\mathbf{x}}_3$. The orthonormal triad $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3$ serves as the basis for an alternative ray-centred coordinate system. The vectors $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ are the two independent and mutually perpendicular shear wave polarizations, satisfying eqs (65), whereas $\hat{\mathbf{e}}_3$ is the unit slowness vector $\hat{\boldsymbol{\sigma}} = \boldsymbol{\sigma}/\|\boldsymbol{\sigma}\|$ along the ray. The designation $\hat{\mathbf{e}}_3$ is our only exception to the convertion that a sans serif font is used to represent inherently 2D (cross-path) quantities.

$$h_1 = 1, \quad h_2 = 1, \quad h_3 = h = 1 - \sigma^{-1}\mathbf{q}\cdot\boldsymbol{\nabla}_\perp\sigma, \tag{66}$$

where $\mathbf{q}$ denotes the perpendicular two-vector eq. (10), and σ and $\boldsymbol{\nabla}_\perp\sigma$ are evaluated on $\boldsymbol{\xi} = (0, 0, l)$, as before. The product $h = h_1h_2h_3$ of the three factors eq. (66) is the Cartesian to ray-centred Jacobian eq. (28). Since the cross-path factors h_1 and h_2 are unity, this 3D Jacobian is simply a unidirectional stretch factor, which accounts for the difference in differential arclengths dl measured along a neighbouring ray $\mathbf{x} = (q_1, q_2, l)$ and the central ray $\boldsymbol{\xi} = (0, 0, l)$.

The conjugate momenta associated with the three ray-centred coordinates q_1, q_2, l are

$$p_1 = \partial T/\partial q_1, \quad p_2 = \partial T/\partial q_2, \quad p_l = \partial T/\partial l. \tag{67}$$

The ray-centred coordinates of the slowness vector at a point $\mathbf{x} = (q_1, q_2, l)$ can be written in terms of these momenta in the form (Farra & Madariaga 1987)

$$\boldsymbol{\sigma} = \left(p_1, p_2, h^{-1}p_l\right). \tag{68}$$

By analogy with the neighbouring-ray offset vector eq. (10) we define the slowness two-vector

$$\mathbf{p} = p_1\hat{\mathbf{e}}_1 + p_2\hat{\mathbf{e}}_2. \tag{69}$$

At a point $\boldsymbol{\xi} = (0, 0, l)$ on the ray the 3D slowness eq. (68) reduces to $\boldsymbol{\sigma} = (0, 0, \sigma)$, as expected.

5.5 Ray-centred Hamiltonian

Upon solving the ray-centred version of the eikonal equation, $\mathbf{p}\cdot\mathbf{p} + h^{-2}p_l^2 = \sigma^2$, for the momentum p_l conjugate to the arclength l, we obtain

$$p_l = h\sqrt{\sigma^2 - \mathbf{p}\cdot\mathbf{p}}. \tag{70}$$

Seeking to eliminate this along-ray momentum variable from further explicit consideration, we introduce a reduced, ray-centred Hamiltonian defined by (Farra & Madariaga 1987)

$$\mathcal{H}(\mathbf{q}, \mathbf{p}, l) = -h\sqrt{\sigma^2 - \mathbf{p}\cdot\mathbf{p}}, \tag{71}$$

where the scalar slowness and scale factor are now considered to be functions of the form $\sigma(\mathbf{q}, l)$ and $h(\mathbf{q}, l)$. The ray-centred version of Hamilton's equations

$$d\mathbf{q}/dl = \partial\mathcal{H}/\partial\mathbf{p}, \qquad d\mathbf{p}/dl = -\partial\mathcal{H}/\partial\mathbf{q} \tag{72}$$

can be integrated to trace the evolution of the cross-path generalized coordinate $\mathbf{q}$ and conjugate momentum $\mathbf{p}$ along a neighbouring ray in a 4D, ray-centred phase space. The reduction of the dimension from six to four is obviously advantageous; however, there is a tradeoff inasmuch as the reduced Hamiltonian $\mathcal{H}(\mathbf{q}, \mathbf{p}, l)$ is no longer conservative, because of the explicit dependence upon the integration variable l. The rate of change of the eliminated momentum variable p_l along a neighbouring ray is

$$dp_l/dl = -d\mathcal{H}/dl = -(\partial\mathcal{H}/\partial\mathbf{q})\cdot(d\mathbf{q}/dl) - (\partial\mathcal{H}/\partial\mathbf{p})\cdot(d\mathbf{p}/dl) - \partial\mathcal{H}/\partial l = -\partial\mathcal{H}/\partial l. \tag{73}$$

The identification eq. (71) of $\mathcal{H}(\mathbf{q}, \mathbf{p}, l)$ with the negative along-ray momentum $-p_l$ is a characteristic feature of this so-called parametric reduction procedure (Lanczos 1962, Chapter 5, Section 10).

5.6 Paraxial approximation

In principle it is possible to use the reduced Hamiltonian formulation in eqs (71) and (72) to trace an arbitrary neighbouring ray described by the offset two-vectors $\mathbf{q}$ and $\mathbf{p}$. In most ray-centred applications, including the present one, however, we are interested only in neighbouring rays that deviate infinitesimally from the central ray. To trace these so-called paraxial rays, we can replace the exact Hamiltonian (71) by its Taylor series expansion about $\mathbf{q} = \mathbf{0}, \mathbf{p} = \mathbf{0}$. Upon carrying out this expansion correct to second order in $\|\mathbf{q}\|$ and $\|\mathbf{p}\|$, we obtain

$$\mathcal{H} = -\sigma + \frac{1}{2}\sigma^{-1}\mathbf{p}\cdot\mathbf{p} - \frac{1}{2}\mathbf{q}\cdot\boldsymbol{\Sigma}\cdot\mathbf{q}, \tag{74}$$

where both the scalar slowness σ and the symmetric two-tensor

$$\boldsymbol{\Sigma} = \boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\sigma - 2\sigma^{-1}(\boldsymbol{\nabla}_\perp\sigma)(\boldsymbol{\nabla}_\perp\sigma) = \boldsymbol{\Sigma}^{\mathrm{T}} \tag{75}$$

are evaluated at points $\boldsymbol{\xi} = (0, 0, l)$ along the central ray. The quadratic dependence of the paraxial Hamiltonian eq. (74) upon $\mathbf{q}$ and $\mathbf{p}$ renders the associated canonical eq. (72) linear in both of these two-vectors:

$$d\mathbf{q}/dl = \sigma^{-1}\mathbf{p}, \qquad d\mathbf{p}/dl = \boldsymbol{\Sigma}\cdot\mathbf{q}. \tag{76}$$

We shall refer to the linear system of eqs (76) as the paraxial ray-tracing equations. To trace a paraxial ray we must integrate these linear equations, subject to the initial conditions

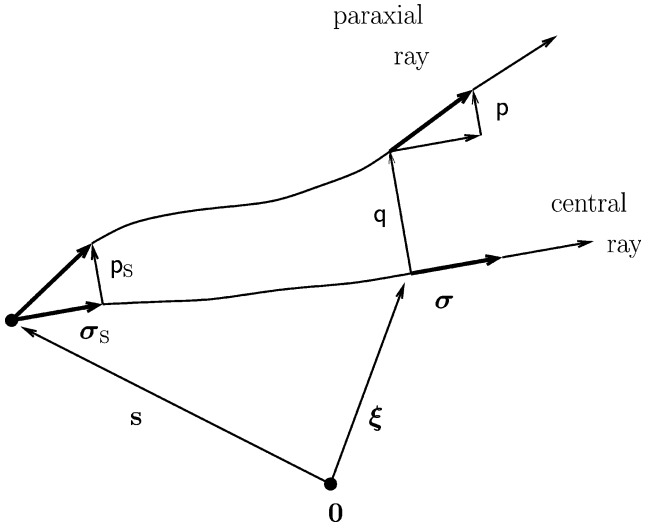

Figure 11. Schematic depiction of a central and a paraxial ray traced from a point source $\mathbf{s}$. The Cartesian three-vectors $\boldsymbol{\xi}$ and $\boldsymbol{\sigma}$ are the position and slowness along the central ray; the initial slowness at the source is denoted by $\boldsymbol{\sigma}_\mathrm{s}$. A paraxial ray is fully described by the paraxial two-vectors $\mathbf{q}$ and $\mathbf{p}$; the first of these specifies the perpendicular offset in position and the second specifies the associated slowness. The vector $\boldsymbol{\sigma}+\mathbf{p}$ is the slowness at the point $\boldsymbol{\xi}+\mathbf{q}$ on the paraxial ray, as shown. The initial value of $\mathbf{p}$ at the source $\mathbf{s}$ is denoted by $\mathbf{p}_\mathrm{s}$.

$$\mathbf{q}(0)=\mathbf{0}, \qquad \mathbf{p}(0)=\mathbf{p}_\mathrm{s}, \tag{77}$$

where the first condition guarantees that the ray departs from the same source location $\mathbf{s}$ and the second specifies the infinitesimal difference in the initial takeoff direction, as illustrated in Fig. 11. The solution in the case $\mathbf{p}_\mathrm{s}=\mathbf{0}$ is $\mathbf{q}=\mathbf{0}$, $\mathbf{p}=\mathbf{0}$ for all $0\le l\le L$; this shows that the paraxial system of eqs (76) and (77) correctly traces the central geometrical ray.

5.7 Traveltime Hessian

The traveltime $T(q_1,q_2,l)$ along a paraxial ray is related to the traveltime $T(0,0,l)$ along the central ray by

$$T(q_1,q_2,l)=T(0,0,l)+\frac{1}{2}\mathbf{q}\cdot\vec{\mathbf{M}}\cdot\mathbf{q}, \tag{78}$$

where the forward-pointing arrow over the symmetric Hessian $\vec{\mathbf{M}}=\vec{\mathbf{M}}^\mathrm{T}$ serves as a reminder that both $T(q_1,q_2,l)$ and $T(0,0,l)$ are measured away from the source point $\mathbf{s}$. Fermat's principle dictates that there is no linear term in the Taylor expansion (78), so the leading term after the zeroth is the quadratic involving the Hessian. Partial differentiation of eq. (78) with respect to q_1, q_2 and l yields the results

$$\mathbf{p}\cdot\mathbf{p}=(\partial T/\partial q_1)^2+(\partial T/\partial q_2)^2=\frac{1}{2}\mathbf{q}\cdot\vec{\mathbf{M}}^2\cdot\mathbf{q}, \tag{79}$$

$$p_l=\partial T/\partial l=\sigma+\frac{1}{2}\mathbf{q}\cdot(d\vec{\mathbf{M}}/dl)\cdot\mathbf{q}. \tag{80}$$

Upon comparing eqs (79) and (80) with the paraxial relation

$$p_l=-\mathcal{H}=\sigma-\frac{1}{2}\sigma^{-1}\mathbf{p}\cdot\mathbf{p}+\frac{1}{2}\mathbf{q}\cdot\boldsymbol{\Sigma}\cdot\mathbf{q}, \tag{81}$$

we obtain a 2×2 Riccati differential equation that governs the evolution of the one-way Hessian:

$$d\vec{\mathbf{M}}/dl+\sigma^{-1}\vec{\mathbf{M}}^2-\boldsymbol{\Sigma}=\mathbf{0}. \tag{82}$$

The backward Hessian $\overleftarrow{\mathbf{M}}$ that measures paraxial traveltime deviations away from the receiver $\mathbf{r}$ satisfies an analogous Riccati equation,

$$d\overleftarrow{\mathbf{M}}/dl-\sigma^{-1}\overleftarrow{\mathbf{M}}^2+\boldsymbol{\Sigma}=\mathbf{0}, \tag{83}$$

obtained by reversing the direction of integration, $dl\to-dl$. Eqs (82) and (83) must be solved subject to the endpoint conditions

$$\vec{\mathbf{M}}\to\sigma_\mathrm{s}l^{-1}\mathbf{I}\ \text{as}\ l\to0, \qquad \overleftarrow{\mathbf{M}}\to\sigma_\mathrm{r}(L-l)^{-1}\mathbf{I}\ \text{as}\ l\to L, \tag{84}$$

where a subscript r denotes evaluation at the receiver $\mathbf{r}$ and $\mathbf{I}$ is the identity two-tensor. We note, finally, that the eigenvalues of $\sigma^{-1}\vec{\mathbf{M}}$ and $\sigma^{-1}\overleftarrow{\mathbf{M}}$ are the principal curvatures of the wave fronts emanating from $\mathbf{s}$ and $\mathbf{r}$, respectively (Červený 1985; Snieder & Chapman 1998).

5.8 Practical integration procedure

Direct numerical integration of eqs (82) and (83) is impractical because of the divergent character of the endpoint conditions (84) at the source **s** and receiver **r**. Focusing upon the forward Hessian $\vec{\mathbf{M}}$ and following Červený (1985) and banana–doughnut I, we seek a solution of the form

$$\vec{\mathbf{M}} = \mathbf{P} \cdot \mathbf{Q}^{-1}. \tag{85}$$

The decomposition eq. (85) converts the forward Riccati eq. (82) into a pair of coupled first-order linear equations:

$$d\mathbf{Q}/dl = \sigma^{-1}\mathbf{P}, \qquad d\mathbf{P}/dl = \mathbf{\Sigma} \cdot \mathbf{Q}. \tag{86}$$

The easily implemented initial conditions at the source **s** equivalent to the first of eq. (84) are

$$\mathbf{Q}(0) = \mathbf{0}, \qquad \mathbf{P}(0) = \mathbf{I}. \tag{87}$$

The identical character of the 2×1 paraxial vector eqs (76) and (77) and the 2×2 eqs (86) and (87) is obviously not accidental. The columns of $\mathbf{Q}$ and $\mathbf{P}$ consist of linearly independent two-vectors $\mathbf{q}$ and $\mathbf{p}$, with associated initial conditions $\mathbf{q}(0) = (0\ 0)^{\mathrm{T}}$, $\mathbf{p}(0) = (1\ 0)^{\mathrm{T}}$ and $\mathbf{q}(0) = (0\ 0)^{\mathrm{T}}$, $\mathbf{p}(0) = (0\ 1)^{\mathrm{T}}$. In a spherically symmetric earth model, the paraxial two-tensors $\mathbf{Q}$ and $\mathbf{P}$ are both diagonal, as discussed in banana–doughnut I. More generally, however, it is not even required that either of these two tensors be symmetric.

5.9 Notational niceties

The paraxial ray-tracing eqs (76) and (86) can be rewritten in terms of the 4×1 and 4×2 column vectors

$$\mathbf{y} = \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}, \qquad \mathbf{Y} = \begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} \tag{88}$$

in the form

$$d\mathbf{y}/dl = \mathbf{A} \cdot \mathbf{y}, \qquad d\mathbf{Y}/dl = \mathbf{A} \cdot \mathbf{Y}, \tag{89}$$

where

$$\mathbf{A} = \begin{pmatrix} \mathbf{0} & \sigma^{-1}\mathbf{I} \\ \mathbf{\Sigma} & \mathbf{0} \end{pmatrix}. \tag{90}$$

The initial conditions equivalent to eqs (77) and (87) are

$$\mathbf{y}(0) = \begin{pmatrix} \mathbf{0} \\ \mathbf{p}_{\mathrm{s}} \end{pmatrix}, \qquad \mathbf{Y}(0) = \begin{pmatrix} \mathbf{0} \\ \mathbf{I} \end{pmatrix}. \tag{91}$$

The ray-centred Hamiltonian (74) can be written in terms of the four-vector $\mathbf{y}$ in the succinct quadratic form

$$\mathcal{H} = -\sigma + \frac{1}{2}\mathbf{y} \cdot \partial^2\mathcal{H}/\partial\mathbf{y}^2 \cdot \mathbf{y}. \tag{92}$$

Following Goldstein (1980; Section 8.1), Červený (1985) and Dahlen & Tromp (1998; Section 15.4.10) we introduce the antisymmetric four-tensor

$$\mathbf{J} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix} = -\mathbf{J}^{\mathrm{T}}. \tag{93}$$

Eqs (89) are then equivalent to

$$d\mathbf{y}/dl = \mathbf{J} \cdot \partial\mathcal{H}/\partial\mathbf{y}, \qquad d\mathbf{Y}/dl = \mathbf{J} \cdot \partial^2\mathcal{H}/\partial\mathbf{y}^2 \cdot \mathbf{Y}. \tag{94}$$

This is the so-called symplectic form of the paraxial ray-tracing equations.

5.10 Forward and backward propagators

The 4×4 propagator $\vec{\mathbf{\Pi}}_{ll'}$ associated with the paraxial eqs (89) is defined by

$$d\vec{\mathbf{\Pi}}_{ll'}/dl = \mathbf{A} \cdot \vec{\mathbf{\Pi}}_{ll'}, \qquad \vec{\mathbf{\Pi}}_{ll} = \mathbf{I}, \tag{95}$$

where $\mathbf{I}$ is the identity four-tensor. The ordered subscripts $0 \le l \le L$ and $0 \le l' \le L$ denote the initial and final points between which paraxial information is propagated along a ray. Note that $\vec{\mathbf{\Pi}}_{ll'}$ is the *forward* propagator from l' to l, even in the case that $l' \ge l$; there is a corresponding *backward* propagator $\overleftarrow{\mathbf{\Pi}}_{ll'}$, defined by

$$d\overleftarrow{\mathbf{\Pi}}_{ll'}/dl = -\mathbf{A}\cdot\overleftarrow{\mathbf{\Pi}}_{ll'}, \qquad \overleftarrow{\mathbf{\Pi}}_{ll} = \mathbf{I}. \tag{96}$$

It is readily verified that the solutions to eqs (95) and (96) satisfy the standard inversion and concatenation properties of a propagator (Gilbert & Backus 1966):

$$\vec{\mathbf{\Pi}}_{l0}\cdot\vec{\mathbf{\Pi}}_{l'0}^{-1} = \vec{\mathbf{\Pi}}_{l0}\cdot\vec{\mathbf{\Pi}}_{0l'} = \vec{\mathbf{\Pi}}_{ll'}, \tag{97}$$

$$\overleftarrow{\mathbf{\Pi}}_{l0}\cdot\overleftarrow{\mathbf{\Pi}}_{l'0}^{-1} = \overleftarrow{\mathbf{\Pi}}_{l0}\cdot\overleftarrow{\mathbf{\Pi}}_{0l'} = \overleftarrow{\mathbf{\Pi}}_{ll'}. \tag{98}$$

Each of the 4×4 propagators $\vec{\mathbf{\Pi}}_{ll'}$ and $\overleftarrow{\mathbf{\Pi}}_{ll'}$ is composed of four 2×2 blocks, which we denote by

$$\vec{\mathbf{\Pi}}_{ll'} = \begin{pmatrix} \tilde{\mathbf{Q}}_{ll'} & \mathbf{Q}_{ll'} \\ \tilde{\mathbf{P}}_{ll'} & \mathbf{P}_{ll'} \end{pmatrix}, \qquad \overleftarrow{\mathbf{\Pi}}_{ll'} = \begin{pmatrix} \tilde{\mathbf{Q}}_{ll'} & -\mathbf{Q}_{ll'} \\ -\tilde{\mathbf{P}}_{ll'} & \mathbf{P}_{ll'} \end{pmatrix}. \tag{99}$$

The two conjugate pairs $\mathbf{Q}_{ll'}$, $\mathbf{P}_{ll'}$ and $\tilde{\mathbf{Q}}_{ll'}$, $\tilde{\mathbf{P}}_{ll'}$ satisfy identical eqs (86) but differing initial conditions,

$$\mathbf{Q}_{ll} = \mathbf{0}, \quad \mathbf{P}_{ll} = \mathbf{I} \quad \text{and} \quad \tilde{\mathbf{Q}}_{ll} = \mathbf{I}, \quad \tilde{\mathbf{P}}_{ll} = \mathbf{0}. \tag{100}$$

The relations (100) obviously guarantee that $\vec{\mathbf{\Pi}}_{ll} = \overleftarrow{\mathbf{\Pi}}_{ll} = \mathbf{I}$. Note that the forward and backward propagators (99) differ only in the signs of the lower-left and upper-right blocks.

5.11 Symplectic symmetry

The symplectic version of the paraxial propagator eqs (76)–(86) is

$$d\vec{\mathbf{\Pi}}_{ll'}/dl = \mathbf{J}\cdot\partial^2\mathcal{H}/\partial\mathbf{y}^2\cdot\vec{\mathbf{\Pi}}_{ll'}, \tag{101}$$

$$d\overleftarrow{\mathbf{\Pi}}_{ll'}/dl = -\mathbf{J}\cdot\partial^2\mathcal{H}/\partial\mathbf{y}^2\cdot\overleftarrow{\mathbf{\Pi}}_{ll'}. \tag{102}$$

It is readily demonstrated from eqs (101) and (102), and the identities $\mathbf{J}\cdot\mathbf{J} = \mathbf{J}^{\mathrm{T}}\cdot\mathbf{J}^{\mathrm{T}} = -\mathbf{I}$ that the forward and backward triple products,

$$\vec{\mathbf{\Pi}}_{ll'}^{\mathrm{T}}\cdot\mathbf{J}\cdot\vec{\mathbf{\Pi}}_{ll'} = \overleftarrow{\mathbf{\Pi}}_{ll'}^{\mathrm{T}}\cdot\mathbf{J}\cdot\overleftarrow{\mathbf{\Pi}}_{ll'} = \mathbf{J}, \tag{103}$$

are invariant along a geometrical ray. Upon taking the determinant of eq. (103) and making use of the initial conditions $\vec{\mathbf{\Pi}}_{ll} = \overleftarrow{\mathbf{\Pi}}_{ll} = \mathbf{I}$, we deduce that $\det\vec{\mathbf{\Pi}}_{ll'} = \det\overleftarrow{\mathbf{\Pi}}_{ll'} = 1$. This result is the ray-centred version of Liouville's theorem.

The symplectic symmetry relations (103) also lead to an explicit representation of the inverse forward and backward propagators. Upon dotting on the left with $\mathbf{J}$ and on the right with $\vec{\mathbf{\Pi}}_{ll'}^{-1}$ and $\overleftarrow{\mathbf{\Pi}}_{ll'}^{-1}$, we obtain

$$\vec{\mathbf{\Pi}}_{ll'}^{-1} = -\mathbf{J}\cdot\vec{\mathbf{\Pi}}_{ll'}^{\mathrm{T}}\cdot\mathbf{J}, \qquad \overleftarrow{\mathbf{\Pi}}_{ll'}^{-1} = -\mathbf{J}\cdot\overleftarrow{\mathbf{\Pi}}_{ll'}^{\mathrm{T}}\cdot\mathbf{J}, \tag{104}$$

or, equivalently,

$$\vec{\mathbf{\Pi}}_{ll'}^{-1} = \begin{pmatrix} \mathbf{P}_{ll'}^{\mathrm{T}} & -\mathbf{Q}_{ll'}^{\mathrm{T}} \\ -\tilde{\mathbf{P}}_{ll'}^{\mathrm{T}} & \tilde{\mathbf{Q}}_{ll'}^{\mathrm{T}} \end{pmatrix}, \qquad \overleftarrow{\mathbf{\Pi}}_{ll'}^{-1} = \begin{pmatrix} \mathbf{P}_{ll'}^{\mathrm{T}} & \mathbf{Q}_{ll'}^{\mathrm{T}} \\ \tilde{\mathbf{P}}_{ll'}^{\mathrm{T}} & \tilde{\mathbf{Q}}_{ll'}^{\mathrm{T}} \end{pmatrix}. \tag{105}$$

Knowing $\vec{\mathbf{\Pi}}_{ll'}$ or $\overleftarrow{\mathbf{\Pi}}_{ll'}$ we can find $\vec{\mathbf{\Pi}}_{ll'}^{-1}$ or $\overleftarrow{\mathbf{\Pi}}_{ll'}^{-1}$ by transposition and rearrangement of the 2×2 blocks, as indicated. Recalling that inversion is equivalent to interchange of the propagator indices, $\vec{\mathbf{\Pi}}_{ll'}^{-1} = \vec{\mathbf{\Pi}}_{l'l}$ and $\overleftarrow{\mathbf{\Pi}}_{ll'}^{-1} = \overleftarrow{\mathbf{\Pi}}_{l'l}$, we see that

$$\mathbf{Q}_{ll'} = -\mathbf{Q}_{l'l}^{\mathrm{T}}, \qquad \mathbf{P}_{ll'} = \tilde{\mathbf{Q}}_{l'l}^{\mathrm{T}}, \tag{106}$$

$$\tilde{\mathbf{Q}}_{ll'} = \mathbf{P}_{l'l}^{\mathrm{T}}, \qquad \tilde{\mathbf{P}}_{ll'} = -\tilde{\mathbf{P}}_{l'l}^{\mathrm{T}}. \tag{107}$$

Expansion of the relations $\vec{\mathbf{\Pi}}_{ll'}\cdot\vec{\mathbf{\Pi}}_{ll'}^{-1} = \vec{\mathbf{\Pi}}_{ll'}^{-1}\cdot\vec{\mathbf{\Pi}}_{ll'} = \mathbf{I}$ and $\overleftarrow{\mathbf{\Pi}}_{ll'}\cdot\overleftarrow{\mathbf{\Pi}}_{ll'}^{-1} = \overleftarrow{\mathbf{\Pi}}_{ll'}^{-1}\cdot\overleftarrow{\mathbf{\Pi}}_{ll'} = \mathbf{I}$ leads to a multitude of 2×2 multiplicative identities, namely

$$\tilde{\mathbf{Q}}_{ll'}\cdot\mathbf{P}_{ll'}^{\mathrm{T}} - \mathbf{Q}_{ll'}\cdot\tilde{\mathbf{P}}_{ll'}^{\mathrm{T}} = \mathbf{P}_{ll'}\cdot\tilde{\mathbf{Q}}_{ll'}^{\mathrm{T}} - \tilde{\mathbf{P}}_{ll'}\cdot\mathbf{Q}_{ll'}^{\mathrm{T}} = \mathbf{I}, \tag{108}$$

$$\mathbf{Q}_{ll'}\cdot\tilde{\mathbf{Q}}_{ll'}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{ll'}\cdot\mathbf{Q}_{ll'}^{\mathrm{T}} = \tilde{\mathbf{P}}_{ll'}\cdot\mathbf{P}_{ll'}^{\mathrm{T}} - \mathbf{P}_{ll'}\cdot\tilde{\mathbf{P}}_{ll'}^{\mathrm{T}} = \mathbf{0}, \tag{109}$$

$$\mathbf{P}_{ll'}^{\mathrm{T}}\cdot\tilde{\mathbf{Q}}_{ll'} - \mathbf{Q}_{ll'}^{\mathrm{T}}\cdot\tilde{\mathbf{P}}_{ll'} = \tilde{\mathbf{Q}}_{ll'}^{\mathrm{T}}\cdot\mathbf{p}_{ll'} - \tilde{\mathbf{P}}_{ll'}^{\mathrm{T}}\cdot\mathbf{Q}_{ll'} = \mathbf{I}, \tag{110}$$

$$\mathbf{P}_{ll'}^{\mathrm{T}}\cdot\mathbf{Q}_{ll'} - \mathbf{Q}_{ll'}^{\mathrm{T}}\cdot\mathbf{P}_{ll'} = \tilde{\mathbf{Q}}_{ll'}^{\mathrm{T}}\cdot\tilde{\mathbf{P}}_{ll'} - \tilde{\mathbf{P}}_{ll'}^{\mathrm{T}}\cdot\tilde{\mathbf{Q}}_{ll'} = \mathbf{0}. \tag{111}$$

Last but not least, we can use eq. (97) and the first of eqs (105) to express each of the l'-to-l elements in terms of the forward propagators $\vec{\mathbf{\Pi}}_{l0}$ and $\vec{\mathbf{\Pi}}_{l'0}$ from the source:

$$\mathbf{Q}_{ll'} = \mathbf{Q}_{l0} \cdot \tilde{\mathbf{Q}}^{\mathrm{T}}_{l'0} - \tilde{\mathbf{Q}}_{l0} \cdot \mathbf{Q}^{\mathrm{T}}_{l'0}, \tag{112}$$

$$\mathbf{P}_{ll'} = \mathbf{P}_{l0} \cdot \tilde{\mathbf{Q}}^{\mathrm{T}}_{l'0} - \tilde{\mathbf{P}}_{l0} \cdot \mathbf{Q}^{\mathrm{T}}_{l'0}, \tag{113}$$

$$\tilde{\mathbf{Q}}_{ll'} = \tilde{\mathbf{Q}}_{l0} \cdot \mathbf{P}^{\mathrm{T}}_{l'0} - \mathbf{Q}_{l0} \cdot \tilde{\mathbf{P}}^{\mathrm{T}}_{l'0}, \tag{114}$$

$$\tilde{\mathbf{P}}_{ll'} = \tilde{\mathbf{P}}_{l0} \cdot \mathbf{P}^{\mathrm{T}}_{l'0} - \mathbf{P}_{l0} \cdot \tilde{\mathbf{P}}^{\mathrm{T}}_{l'0}. \tag{115}$$

Using the relations (112)–(115) we can compute all of the entries (99) and (105) of both 4×4 propagators $\vec{\mathbf{\Pi}}_{ll'}$, $\overleftarrow{\mathbf{\Pi}}_{ll'}$ and their inverses $\vec{\mathbf{\Pi}}^{-1}_{ll'}$, $\overleftarrow{\mathbf{\Pi}}^{-1}_{ll'}$ with a single forward integration that starts at the source $\mathbf{s}$ and ends at the receiver $\mathbf{r}$.

5.12 Geometrical spreading

The geometrical spreading of infinite-frequency waves about the central ray is governed by the point-source Jacobian, which we take to be defined by (Červený 1985)

$$J = \det \mathbf{Q}_{l0}. \tag{116}$$

The signed determinant (116) starts out singularly at the source,

$$J \to \sigma_{\mathrm{s}}^{-2} l^2 \quad \text{as} \quad l \to 0, \tag{117}$$

and varies smoothly along a ray, passing through zero at every caustic. An alternative measure of geometrical spreading is provided by the inherently positive quantity analogous to the straight-ray source–receiver distance $\|\mathbf{r} - \mathbf{s}\|$ in a homogeneous medium (Aki & Richards 1980, Section 4.4.3):

$$\mathcal{R} = \sqrt{|d\Sigma|/d\Omega}, \tag{118}$$

where $d\Sigma$ is the differential cross-sectional area of a ray tube at a point $\boldsymbol{\xi}$ and $d\Omega$ is the solid angle subtended by the tube at the source point $\mathbf{s}$, as illustrated in Fig. 12. the geometrical spreading factor (118) of Aki & Richards and the Jacobian (116) of Červený are related everywhere by

$$\mathcal{R} = \sigma_{\mathrm{s}} \sqrt{|J|}. \tag{119}$$

The limiting behaviour of the spreading factor near the source $\mathbf{s}$ is $\mathcal{R} \to l$ as $l \to 0$, as expected.

5.13 Acoustic response

For the purpose of illustration, we shall consider a specific example—an isotropic, explosive source in an acoustic medium. The frequency-domain response to such a source is given by (Morse & Ingard 1968, Section 7.1; banana–doughnut II)

$$u(\omega) = \sum_{\text{rays}} A \dot{m}(\omega) \exp i(-\omega T + M\pi/2), \tag{120}$$

where $u(\omega)$ denotes (somewhat unconventionally) the incremental pressure, and where we have let

$$A = \frac{1}{4\pi} (\rho_{\mathrm{s}} \rho_{\mathrm{r}})^{1/2} (\sigma_{\mathrm{s}} \sigma_{\mathrm{r}})^{-1/2} |J|^{-1/2}. \tag{121}$$

The frequency-dependent term $\dot{m}(\omega)$ is the Fourier transform of the second time derivative of an infinitesimally small volume situated at the source, $\dot{m}(t) = d\ddot{V}(t)$, and the phase factors T and M are the traveltime and Maslov index, or the number of caustic passages, between the

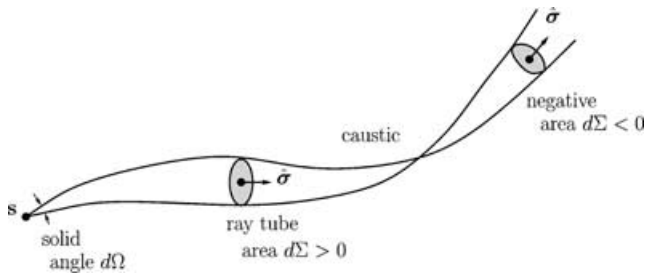

Figure 12. Schematic illustration of a ray tube subtending an infinitesimal solid angle $d\Omega$ at the source $\mathbf{s}$. The differential cross-sectional area $d\Sigma$ varies along the central geometrical ray. The area is positive, $d\Sigma > 0$, until the first passage through a caustic, after which it may be informally regarded as negative, $d\Sigma < 0$. The sketch is somewhat misleading since in passing through a caustic, a ray tube closes in only one direction rather than two.

source **s** and receiver **r**, respectively. The quantity ρ in the acoustic amplitude factor (121) is the mass density, and the subscripts s and r denote evaluation at the source **s** and receiver **r**, as above. The summation accounts for multipathing, that is, more than one geometrical ray between **s** and **r**.

5.14 Source–receiver reciprocity

Upon specializing the first of the identities (106) to the case of the source-to-receiver subpropagator, we see that the Jacobian (116) satisfies

$$|J| = |\det \mathbf{Q}_{L0}| = |\det \mathbf{Q}_{0L}|. \tag{122}$$

Eq. (122) guarantees that the ray-theoretical amplitude (121) of a body wave is invariant under an interchange $\mathbf{s} \leftrightarrow \mathbf{r}$ of the source and receiver (Snieder & Chapman 1998):

$$A_{\rm rs} = A_{\rm sr}. \tag{123}$$

The traveltime and Maslov index satisfy similar relations, $T_{\rm rs} = T_{\rm sr}$ and $M_{\rm rs} = M_{\rm sr}$, so that the unperturbed ray-theoretical response (120) is consistent with the principle of source–receiver reciprocity: $u_{\rm rs}(\omega) = u_{\rm sr}(\omega)$.

6 RAY PERTURBATION THEORY

Suppose now that the density and acoustic slowness are subjected to an infinitesimal perturbation:

$$\rho \to \rho + \delta\rho, \qquad \sigma \to \sigma + \delta\sigma. \tag{124}$$

The resulting perturbation in the ray-centred Hamiltonian,

$$\mathcal{H} \to \mathcal{H} + \delta\mathcal{H}, \tag{125}$$

will give rise to a perturbation in the ray path and geometrical spreading between the source and receiver,

$$\mathbf{q} \to \mathbf{q} + \delta\mathbf{q}, \qquad \mathbf{p} \to \mathbf{p} + \delta\mathbf{p}, \qquad \mathbf{y} \to \mathbf{y} + \delta\mathbf{y}, \tag{126}$$

$$\mathbf{Q} \to \mathbf{Q} + \delta\mathbf{Q}, \qquad \mathbf{P} \to \mathbf{P} + \delta\mathbf{P}, \qquad \mathbf{Y} \to \mathbf{Y} + \delta\mathbf{Y}, \tag{127}$$

as well as to the arrival time and amplitude of the pressure pulse (120):

$$T \to T + \delta T, \qquad A \to A + \delta A. \tag{128}$$

It is well known that the first-order traveltime perturbation δT is given by the 1D integral (34) of the slowness perturbation $\delta\sigma$ along the unperturbed ray. We focus attention upon the ray-theoretical amplitude perturbation δA in the remainder of this section.

6.1 Perturbed Hamiltonian

The first-order perturbation $\delta\mathcal{H}(\mathbf{q}, \mathbf{p}, l)$ to the Hamiltonian (71) along a fixed path $\mathbf{q}, \mathbf{p}$ is given by (Farra & Madariaga 1987)

$$\delta\mathcal{H} = \delta\sigma(\partial\mathcal{H}/\partial\sigma) = -\delta\sigma\, h\sigma(\sigma^2 - \mathbf{p}\cdot\mathbf{p})^{-1/2}. \tag{129}$$

There is no term involving δh in accordance with the stipulation that the coordinate system is unaffected by a change in the earth model (Nowack & Lutter 1988). Upon expanding the right-hand side of eq. (129) correct to second order in $\|\mathbf{q}\|$ and $\|\mathbf{p}\|$, we find that

$$\delta\mathcal{H} = -\delta\sigma - \mathbf{q}\cdot\left(\boldsymbol{\nabla}_\perp\delta\sigma - \sigma^{-1}\delta\sigma\boldsymbol{\nabla}_\perp\sigma\right) - \frac{1}{2}\sigma^{-2}\delta\sigma\mathbf{p}\cdot\mathbf{p} - \frac{1}{2}\mathbf{q}\cdot\left(\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\delta\sigma - 2\sigma^{-1}\boldsymbol{\nabla}_\perp\sigma\boldsymbol{\nabla}_\perp\delta\sigma\right)\cdot\mathbf{q}, \tag{130}$$

where the quantities σ, $\delta\sigma$ and all of their cross-path gradients are evaluated on the unperturbed geometrical ray $\boldsymbol{\xi} = (0, 0, l)$.

6.2 Ray path perturbation

The perturbation $\delta\mathbf{q}$, $\delta\mathbf{p}$ to the ray path is found by perturbing the first of (94):

$$d(\delta\mathbf{y})/dl = \mathbf{J}\cdot\partial^2\mathcal{H}/\partial\mathbf{y}^2\cdot\delta\mathbf{y} + \mathbf{J}\cdot\partial(\delta\mathcal{H})/\partial\mathbf{y}, \tag{131}$$

where the partial derivatives on the right are all evaluated on the unperturbed ray $\mathbf{y} = \mathbf{0}$. Eq. (131) is an inhomogeneous version of the first of eqs (89), with a 'source' term $\mathbf{f} = \mathbf{J}\cdot\partial(\delta\mathcal{H})/\partial\mathbf{y}$ arising from the perturbation (129) to the Hamiltonian (Nowack & Lutter 1988; Snieder & Sambridge 1992):

$$d(\delta\mathbf{y})/dl = \mathbf{A}\cdot\delta\mathbf{y} + \mathbf{f}, \tag{132}$$

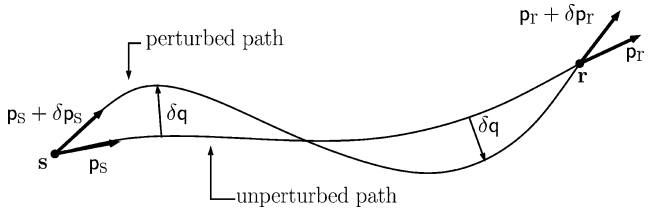

Figure 13. The unperturbed ray in the unperturbed medium σ and the perturbed ray defined by its perpendicular offset vector $\delta\mathbf{q}$ in the perturbed medium $\sigma + \delta\sigma$ both pass through the source $\mathbf{s}$ and the receiver $\mathbf{r}$. The initial and final slowness vectors are perturbed by amounts $\mathbf{p}_s \to \mathbf{p}_s + \delta\mathbf{p}_s$ and $\mathbf{p}_r \to \mathbf{p}_r + \delta\mathbf{p}_r$.

where

$$\mathbf{f} = \begin{pmatrix} \mathbf{0} \\ \boldsymbol{\nabla}_\perp \delta\sigma - \sigma^{-1}\delta\sigma\boldsymbol{\nabla}_\perp\sigma \end{pmatrix}. \tag{133}$$

The solution to eq. (132) can be expressed in terms of the forward propagator $\vec{\boldsymbol{\Pi}}_{ll'}$ in the form

$$\delta\mathbf{y}(l) = \int_0^l \vec{\boldsymbol{\Pi}}_{ll'} \cdot \mathbf{f}(l')\, dl' + \vec{\boldsymbol{\Pi}}_{l0} \cdot \delta\mathbf{y}(0). \tag{134}$$

The endpoint conditions guaranteeing that the perturbed ray passes through both the source $\mathbf{s}$ and the receiver $\mathbf{r}$ are (see Fig. 13)

$$\delta\mathbf{y}(0) = \begin{pmatrix} \mathbf{0} \\ \delta\mathbf{p}_s \end{pmatrix}, \qquad \delta\mathbf{y}(L) = \begin{pmatrix} \mathbf{0} \\ \delta\mathbf{p}_r \end{pmatrix}. \tag{135}$$

The perturbations in the initial and final slownesses $\delta\mathbf{p}_s$, $\delta\mathbf{p}_r$ are found from eq. (134) to be

$$\delta\mathbf{p}_s = -\mathbf{Q}_{L0}^{-1} \cdot \int_0^L \mathbf{Q}_{Ll} \cdot \left(\boldsymbol{\nabla}_\perp \delta\sigma_l - \sigma_l^{-1}\delta\sigma_l \boldsymbol{\nabla}_\perp \sigma_l \right) dl, \tag{136}$$

$$\delta\mathbf{p}_r = \mathbf{Q}_{L0}^{-\mathrm{T}} \cdot \int_0^L \mathbf{Q}_{l0}^{\mathrm{T}} \cdot \left(\boldsymbol{\nabla}_\perp \delta\sigma_l - \sigma_l^{-1}\delta\sigma_l \boldsymbol{\nabla}_\perp \sigma_l \right) dl, \tag{137}$$

where the subscripts on σ_l, $\delta\sigma_l$ and their cross-path derivatives denote evaluation at $0 \le l \le L$, and where we have used eqs (112) and (113) and the as yet underived relation (155) to reduce the latter expression. The perturbation in the pointwise position $\delta\mathbf{q}_l = \delta\mathbf{q}(l)$ of the path is given by

$$\delta\mathbf{q}_l = \int_0^L \mathbf{G}_{ll'} \cdot \left(\boldsymbol{\nabla}_\perp \delta\sigma_{l'} - \sigma_{l'}^{-1}\delta\sigma_{l'} \boldsymbol{\nabla}_\perp \sigma_{l'} \right) dl', \tag{138}$$

where

$$\mathbf{G}_{ll'} = \begin{cases} \mathbf{Q}_{l0} \cdot \mathbf{Q}_{L0}^{-1} \cdot \tilde{\mathbf{Q}}_{L0} \cdot \mathbf{Q}_{l'0}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{l0} \cdot \mathbf{Q}_{l'0}^{\mathrm{T}} & \text{if} \quad l' \le l \\ \mathbf{Q}_{l0} \cdot \mathbf{Q}_{L0}^{-1} \cdot \tilde{\mathbf{Q}}_{L0} \cdot \mathbf{Q}_{l'0}^{\mathrm{T}} - \mathbf{Q}_{l0} \cdot \tilde{\mathbf{Q}}_{l'0}^{\mathrm{T}} & \text{if} \quad l' \ge l. \end{cases} \tag{139}$$

The quantity $\mathbf{G}_{ll'}$ is the 2×2 Green tensor for two-point paraxial ray tracing; it is readily verified that this tensor satisfies the source–receiver reciprocity relation $\mathbf{G}_{ll'} = \mathbf{G}_{l'l}^{\mathrm{T}}$.

6.3 Amplitude perturbation

The fractional change in amplitude $\delta(\ln A) = \delta A / A$ is given, correct to first order in the perturbations (124) by

$$\delta(\ln A) = \frac{1}{2}\left(\frac{\delta\rho_s}{\rho_s} + \frac{\delta\rho_r}{\rho_r} - \frac{\delta\sigma_s}{\sigma_s} - \frac{\delta\sigma_r}{\sigma_r} \right) - \frac{1}{2}\frac{\delta|J|}{|J|}. \tag{140}$$

The initial term arising from the change in the properties of the medium at the source $\mathbf{s}$ and receiver $\mathbf{r}$ pertains specifically to the case (121) of an isotropic source in an acoustic medium. In the seismologically relevant case of a moment-tensor source in an elastic earth model, this endpoint contribution is replaced by a more complicated expression that depends, among other things, upon the perturbations $\delta\mathbf{p}_s$ and $\delta\mathbf{p}_r$ in the initial and final slownesses of the ray at $\mathbf{s}$ and $\mathbf{r}$ (see, e.g. Dahlen & Tromp 1998 Section 15.9.5). We shall not be concerned with the endpoint contributions to $\delta(\ln A)$ in the remainder of this paper; our sole emphasis is upon the final term in eq. (140), which expresses the contribution owing to 3D focusing and defocusing of the ray tube between $\mathbf{s}$ and $\mathbf{r}$. Reserving the designation $\delta(\ln A)$ for this term alone, we shall henceforth restrict attention to the quantity

$$\delta(\ln A) = -\frac{1}{2}\frac{\delta|J|}{|J|} = -\frac{1}{2}\frac{\delta(\det \mathbf{Q}_{0L})}{\det \mathbf{Q}_{0L}}, \tag{141}$$

where the latter equality is valid regardless of the sign of $J = \det \mathbf{Q}_{0L}$. This strictly geometrical relation is generic; that is, it describes the first-order perturbation in the logarithmic amplitude owing to the perturbation in the absolute point-source Jacobian $|J|$ in an arbitrary smooth acoustic or isotropic elastic medium.

6.4 Only known seismological application of the Cayley–Hamilton theorem?

The perturbation of the determinant in eq. (141) is readily shown by expansion of an arbitrary perturbed two-tensor $\delta\mathbf{Q}$ to be given by

$$\delta(\det \mathbf{Q}) = (\operatorname{tr} \mathbf{Q})(\operatorname{tr} \delta\mathbf{Q}) - \operatorname{tr}(\mathbf{Q} \cdot \delta\mathbf{Q}), \tag{142}$$

where tr denotes the trace. The inverse of $\mathbf{Q}$ can be written with the aid of the Cayley–Hamilton theorem (Horn & Johnson 1985, Section 2.4),

$$\mathbf{Q}^2 - (\operatorname{tr} \mathbf{Q})\,\mathbf{Q} + (\det \mathbf{Q})\,\mathbf{I} = \mathbf{0}, \tag{143}$$

in the form

$$\mathbf{Q}^{-1} = \frac{(\operatorname{tr} \mathbf{Q})\,\mathbf{I} - \mathbf{Q}}{\det \mathbf{Q}}. \tag{144}$$

Upon making use of eqs (142) and (144) we find that the logarithmic amplitude perturbation (141) reduces to

$$\delta(\ln A) = -\frac{1}{2} \operatorname{tr}\left(\mathbf{Q}_{L0}^{-1} \cdot \delta\mathbf{Q}_{L0}\right). \tag{145}$$

Eq. (145) is the most convenient form of $\delta(\ln A)$ for our purposes. We now turn to the crux of the problem—the determination of the perturbed source-to-receiver subpropagator $\delta\mathbf{Q}_{L0}$.

6.5 Perturbation in the geometrical spreading

Upon perturbing the second of eqs (94) we obtain, by analogy with (131),

$$d(\delta\mathbf{Y})/dl = \mathbf{J} \cdot \partial^2\mathcal{H}/\partial\mathbf{y}^2 \cdot \delta\mathbf{Y} + \mathbf{J} \cdot [\partial^2(\delta\mathcal{H})/\partial\mathbf{y}^2 + \delta\mathbf{y} \cdot \partial^3\mathcal{H}/\partial\mathbf{y}^3] \cdot \mathbf{Y}, \tag{146}$$

where, as usual, the partial derivatives are all evaluated on the unperturbed ray $\mathbf{y} = \mathbf{0}$. Eq. (146) is an inhomogeneous version of the second of eqs (89), of the form

$$d(\delta\mathbf{Y})/dl = \mathbf{A} \cdot \delta\mathbf{Y} + \delta\mathbf{A} \cdot \mathbf{Y}. \tag{147}$$

The associated initial condition obtained by perturbing the second of eqs (91) is

$$\delta\mathbf{Y}(0) = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}. \tag{148}$$

In this case there are two contributions to the 'source' term $\delta\mathbf{A} \cdot \mathbf{Y}$, the first arising from the perturbation to the Hamiltonian $\delta\mathcal{H}$ and the second owing to the associated perturbation $\delta\mathbf{y}$ in the ray path. The latter contribution accounts for the fact that the desired quantity $\mathbf{Y} + \delta\mathbf{Y}$ is a measure of the spreading about the perturbed ray path $\mathbf{y} + \delta\mathbf{y}$ (Farra & Madariaga 1987; Nowack & Lutter 1988; Coates & Chapman 1990). The quantity $\partial^2(\delta\mathcal{H})/\partial\mathbf{y}^2$ is easily found by differentiation of the Hamiltonian perturbation (130); to find $\delta\mathbf{y} \cdot \partial^3\mathcal{H}/\partial\mathbf{y}^3$, it is necessary to extend the expansion (74) of the unperturbed Hamiltonian eq. (71) to third order in $\|\mathbf{q}\|$ and $\|\mathbf{p}\|$. Upon carrying out this calculation, we obtain

$$\delta\mathbf{A} = \begin{pmatrix} \mathbf{0} & -\sigma^{-2}\delta\sigma\mathbf{I} \\ \delta\boldsymbol{\Sigma} & \mathbf{0} \end{pmatrix}, \tag{149}$$

where

$$\begin{aligned}\delta\boldsymbol{\Sigma} = {} & \boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\delta\sigma - \sigma^{-1}(\boldsymbol{\nabla}_\perp\sigma\boldsymbol{\nabla}_\perp\delta\sigma + \boldsymbol{\nabla}_\perp\delta\sigma\boldsymbol{\nabla}_\perp\sigma) + \delta\mathbf{q} \cdot \boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\sigma - \sigma^{-1}(\delta\mathbf{q} \cdot \boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\sigma)\boldsymbol{\nabla}_\perp\sigma - \sigma^{-1}\boldsymbol{\nabla}_\perp\sigma(\delta\mathbf{q} \cdot \boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\sigma) \\ & -\sigma^{-1}(\delta\mathbf{q} \cdot \boldsymbol{\nabla}_\perp\sigma)\left(\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\sigma + 3\sigma^{-1}\boldsymbol{\nabla}_\perp\sigma\boldsymbol{\nabla}_\perp\sigma\right) = \delta\boldsymbol{\Sigma}^{\mathrm{T}}.\end{aligned} \tag{150}$$

The terms that are explicitly dependent upon $\delta\sigma$ and its cross-path gradients $\boldsymbol{\nabla}_\perp\delta\sigma$ and $\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\delta\sigma$ arise from the perturbation to the Hamiltonian (130), whereas the terms dependent upon the two vector $\delta\mathbf{q}$ arise from the path perturbation (138). The solution to eqs (147) and (148) is

$$\delta\mathbf{Y}(l) = \int_0^l \vec{\boldsymbol{\Pi}}_{ll'} \cdot \delta\mathbf{A}(l') \cdot \mathbf{Y}(l')\, dl'. \tag{151}$$

In particular, the desired 2×2 perturbation $\delta\mathbf{Q}_{L0}$ at the receiver $\mathbf{r}$ is given by

$$\delta\mathbf{Q}_{L0} = \int_0^L \sigma_l^{-2}\delta\sigma_l\left(\mathbf{Q}_{L0} \cdot \tilde{\mathbf{P}}_{l0}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{L0} \cdot \mathbf{P}_{l0}^{\mathrm{T}}\right) \cdot \mathbf{P}_{l0}\, dl + \int_0^L \left(\mathbf{Q}_{L0} \cdot \tilde{\mathbf{Q}}_{l0}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{L0} \cdot \mathbf{Q}_{l0}^{\mathrm{T}}\right) \cdot \delta\boldsymbol{\Sigma}_l \cdot \mathbf{Q}_{l0}\, dl, \tag{152}$$

where the subscripts upon σ_l, $\delta\sigma_l$ and $\delta\mathbf{\Sigma}_l$ denote evaluation at $0 \le l \le L$, and where we have used eqs (112) and (114) to express the result in terms of $\mathbf{Q}_{l0}$, $\mathbf{P}_{l0}$ and $\tilde{\mathbf{Q}}_{l0}$, $\tilde{\mathbf{P}}_{l0}$.

6.6 Two useful Hessian formulae

The two-way traveltime Hessian (12) is given in terms of the 2×2 subpropagators by

$$\overleftrightarrow{\mathbf{M}} = \vec{\mathbf{M}} + \overleftarrow{\mathbf{M}} = \mathbf{P}_{l0} \cdot \mathbf{Q}_{l0}^{-1} - \mathbf{P}_{lL} \cdot \mathbf{Q}_{lL}^{-1}, \tag{153}$$

where the minus sign in front of the second term is present in accordance with eqs (99), inasmuch as $\overleftarrow{\mathbf{M}}$ is the backward Hessian. Upon utilizing eqs (112) and (113) in (153) we obtain

$$\begin{aligned}\overleftrightarrow{\mathbf{M}} &= \mathbf{P}_{l0} \cdot \mathbf{Q}_{l0}^{-1} - \left(\mathbf{P}_{l0} \cdot \tilde{\mathbf{Q}}_{L0}^{\mathrm{T}} - \tilde{\mathbf{P}}_{l0} \cdot \mathbf{Q}_{L0}^{\mathrm{T}}\right) \cdot \left(\mathbf{Q}_{l0} \cdot \tilde{\mathbf{Q}}_{L0}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{l0} \cdot \mathbf{Q}_{L0}^{\mathrm{T}}\right)^{-1} \\ &= \left(\tilde{\mathbf{P}}_{l0} \cdot \mathbf{Q}_{L0}^{\mathrm{T}} - \mathbf{P}_{l0} \cdot \mathbf{Q}_{l0}^{-1} \cdot \tilde{\mathbf{Q}}_{l0} \cdot \mathbf{Q}_{L0}^{\mathrm{T}}\right) \cdot \left(\mathbf{Q}_{l0} \cdot \tilde{\mathbf{Q}}_{L0}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{l0} \cdot \mathbf{Q}_{L0}^{\mathrm{T}}\right)^{-1}.\end{aligned} \tag{154}$$

The triple dot product $\mathbf{P}_{l0} \cdot \mathbf{Q}_{l0}^{-1} \cdot \tilde{\mathbf{Q}}_{l0}$ in the final form of eq. (154) can be rearranged with the aid of the identities (108) and (109):

$$\mathbf{P}_{l0} \cdot \mathbf{Q}_{l0}^{-1} \cdot \tilde{\mathbf{Q}}_{l0} = \mathbf{P}_{l0} \cdot \mathbf{Q}_{l0}^{-1} \cdot \tilde{\mathbf{Q}}_{l0} \cdot \mathbf{Q}_{l0}^{\mathrm{T}} \cdot \mathbf{Q}_{l0}^{-\mathrm{T}} = \mathbf{P}_{l0} \cdot \mathbf{Q}_{l0}^{-1} \cdot \mathbf{Q}_{l0} \cdot \tilde{\mathbf{Q}}_{l0}^{\mathrm{T}} \cdot \mathbf{Q}_{l0}^{-\mathrm{T}} = \mathbf{P}_{l0} \cdot \tilde{\mathbf{Q}}_{l0}^{\mathrm{T}} \cdot \mathbf{Q}_{l0}^{-\mathrm{T}} \left(\mathbf{I} + \tilde{\mathbf{P}}_{l0} \cdot \mathbf{Q}_{l0}^{\mathrm{T}}\right) \cdot \mathbf{Q}_{l0}^{-\mathrm{T}} = \mathbf{Q}_{l0}^{-\mathrm{T}} + \tilde{\mathbf{P}}_{l0}. \tag{155}$$

Substitution of eq. (155) into (154) leads to the most convenient representation of the Hessian for our purposes:

$$\overleftrightarrow{\mathbf{M}} = \mathbf{Q}_{l0}^{-\mathrm{T}} \cdot \mathbf{Q}_{L0}^{\mathrm{T}} \cdot \left(\tilde{\mathbf{Q}}_{l0} \cdot \mathbf{Q}_{L0}^{\mathrm{T}} - \mathbf{Q}_{l0} \cdot \tilde{\mathbf{Q}}_{L0}^{\mathrm{T}}\right)^{-1}. \tag{156}$$

The quantity we really require is the inverse of this

$$\overleftrightarrow{\mathbf{M}}^{-1} = \left(\tilde{\mathbf{Q}}_{l0} \cdot \mathbf{Q}_{L0}^{\mathrm{T}} - \mathbf{Q}_{l0} \cdot \tilde{\mathbf{Q}}_{L0}^{\mathrm{T}}\right) \cdot \mathbf{Q}_{L0}^{-\mathrm{T}} \cdot \mathbf{Q}_{l0}^{\mathrm{T}} = \mathbf{Q}_{l0} \cdot \mathbf{Q}_{L0}^{-1} \cdot \left(\mathbf{Q}_{L0} \cdot \tilde{\mathbf{Q}}_{l0}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{L0} \cdot \mathbf{Q}_{l0}^{\mathrm{T}}\right), \tag{157}$$

where the final more useful expression is a consequence of the symmetry $\overleftrightarrow{\mathbf{M}}^{-1} = \overleftrightarrow{\mathbf{M}}^{-\mathrm{T}}$.

The transpose of the first two lines of eq. (154) can be rearranged to yield

$$\mathbf{Q}_{L0} \cdot \tilde{\mathbf{P}}_{l0}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{L0} \cdot \mathbf{P}_{l0}^{\mathrm{T}} = -\left(\mathbf{Q}_{L0} \cdot \tilde{\mathbf{Q}}_{l0}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{L0} \cdot \mathbf{Q}_{l0}^{\mathrm{T}}\right) \cdot \overleftarrow{\mathbf{M}}, \tag{158}$$

where $\overleftarrow{\mathbf{M}} = \overleftrightarrow{\mathbf{M}} - \mathbf{P}_{l0} \cdot \mathbf{Q}_{l0}^{-1} = \overleftarrow{\mathbf{M}}^{\mathrm{T}}$ is the backward Hessian, measured from the receiver $\mathbf{r}$. Multiplication of this result on the left by $\mathbf{P}_{l0} \cdot \mathbf{Q}_{L0}^{-1}$ yields

$$\begin{aligned}\mathbf{P}_{l0} \cdot \mathbf{Q}_{L0}^{-1} \cdot \left(\mathbf{Q}_{L0} \cdot \tilde{\mathbf{P}}_{l0}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{L0} \cdot \mathbf{P}_{l0}^{\mathrm{T}}\right) &= -\mathbf{P}_{l0} \cdot \mathbf{Q}_{L0}^{-1} \cdot \left(\mathbf{Q}_{L0} \cdot \tilde{\mathbf{Q}}_{l0}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{L0} \cdot \mathbf{Q}_{l0}^{\mathrm{T}}\right) \cdot \overleftarrow{\mathbf{M}} \\ &= -\left[\mathbf{P}_{l0} \cdot \mathbf{Q}_{l0}^{-1}\right] \cdot \left[\mathbf{Q}_{l0} \cdot \mathbf{Q}_{L0}^{-1} \cdot \left(\mathbf{Q}_{L0} \cdot \tilde{\mathbf{Q}}_{l0}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{L0} \cdot \mathbf{Q}_{l0}^{\mathrm{T}}\right)\right] \cdot \overleftarrow{\mathbf{M}}.\end{aligned} \tag{159}$$

Upon grouping terms as indicated by the square brackets, we obtain a second useful identity:

$$-\vec{\mathbf{M}} \cdot \overleftrightarrow{\mathbf{M}}^{-1} \cdot \overleftarrow{\mathbf{M}} = \mathbf{P}_{l0} \cdot \mathbf{Q}_{L0}^{-1} \cdot \left(\mathbf{Q}_{L0} \cdot \tilde{\mathbf{P}}_{l0}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{L0} \cdot \mathbf{P}_{l0}^{\mathrm{T}}\right). \tag{160}$$

It is noteworthy that the expressions on the right-hand sides of eqs (157) and (160) are identical, except for the substitutions $\mathbf{Q}_{l0} \leftrightarrow \mathbf{P}_{l0}$ and $\tilde{\mathbf{Q}}_{l0} \leftrightarrow \tilde{\mathbf{P}}_{l0}$.

6.7 The home stretch

Upon inserting the result (152) into (145) and rearranging terms we obtain

$$\delta(\ln A) = -\frac{1}{2}\int_0^L \sigma_l^{-2}\delta\sigma_l \,\mathrm{tr}\left[\mathbf{P}_{l0} \cdot \mathbf{Q}_{L0}^{-1} \cdot \left(\mathbf{Q}_{L0} \cdot \tilde{\mathbf{P}}_{l0}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{L0} \cdot \mathbf{P}_{l0}^{\mathrm{T}}\right)\right] dl - \frac{1}{2}\int_0^L \left[\mathbf{Q}_{l0} \cdot \mathbf{Q}_{L0}^{-1} \cdot \left(\mathbf{Q}_{L0} \cdot \tilde{\mathbf{Q}}_{l0}^{\mathrm{T}} - \tilde{\mathbf{Q}}_{L0} \cdot \mathbf{q}_{l0}^{\mathrm{T}}\right)\right] : \delta\mathbf{\Sigma}_l \, dl. \tag{161}$$

The expressions in square brackets are precisely those we laboriously simplified in eqs (157) and (160). Upon utilizing these results, we obtain the desired ray-theoretical representation of the logarithmic amplitude perturbation:

$$\delta(\ln A) = \frac{1}{2}\int_0^L [\sigma^{-2}\delta\sigma \,\mathrm{tr}(\vec{\mathbf{M}} \cdot \overleftrightarrow{\mathbf{M}}^{-1} \cdot \overleftarrow{\mathbf{M}}) - \overleftrightarrow{\mathbf{M}}^{-1} : \delta\mathbf{\Sigma}] \, dl,$$

where we have dropped the subscripts upon σ_l, $\delta\sigma_l$ and $\delta\mathbf{\Sigma}_l$ for simplicity. The full dependence upon the slowness perturbation $\delta\sigma$ and its cross-path gradients can be deduced by substituting the representation (150) of the 2×2 perturbation $\delta\mathbf{\Sigma} = \delta\mathbf{\Sigma}^{\mathrm{T}}$.

6.8 Ray-theoretical amplitude perturbation

Ignoring endpoint contributions such as those in eq. (140), the first-order perturbation to the ray-theoretical amplitude of a body wave is

$$\delta(\ln A) = \frac{1}{2}\int_0^L \sigma^{-2}\delta\sigma\,\mathrm{tr}(\vec{\mathbf{M}}\cdot\overleftrightarrow{\mathbf{M}}^{-1}\cdot\overleftarrow{\mathbf{M}})\,dl - \frac{1}{2}\int_0^L \overleftrightarrow{\mathbf{M}}^{-1} : \Big(\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\delta\sigma - 2\sigma^{-1}\boldsymbol{\nabla}_\perp\sigma\boldsymbol{\nabla}_\perp\delta\sigma\Big)\,dl$$
$$-\frac{1}{2}\int_0^L \overleftrightarrow{\mathbf{M}}^{-1} : \Big[\delta\mathbf{q}\cdot\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\sigma - 2\sigma^{-1}(\delta\mathbf{q}\cdot\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\sigma)\boldsymbol{\nabla}_\perp\sigma - \sigma^{-1}(\delta\mathbf{q}\cdot\boldsymbol{\nabla}_\perp\sigma)\Big(\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\sigma + 3\sigma^{-1}\boldsymbol{\nabla}_\perp\sigma\boldsymbol{\nabla}_\perp\sigma\Big)\Big]\,dl. \tag{163}$$

where the first two integrals represent the direct effect of the Hamiltonian perturbation $\delta\mathcal{H}$, and the final integral is a result of the path perturbation $\delta\mathbf{q}$.

7 DISCUSSION

Before comparing the above result with the $\omega\to\infty$ Born approximation (53) we wish to make a few remarks from a strictly ray-theoretical perspective. We note first that the explicit formula (163) is not some sort of 'silver bullet' that enables ray-theoretical amplitude perturbations $\delta(\ln A)$ to be computed more efficiently than any of the numerical recipes described by Thomson (1983), Nowack & Lutter (1988), Neele *et al.* (1993) or Liu & Tromp (1996). The forward, backward and two-way traveltime Hessians $\vec{\mathbf{M}}$, $\overleftarrow{\mathbf{M}}$ and $\overleftrightarrow{\mathbf{M}}$ must still be computed at every point $0\le l\le L$ along the ray by numerical integration of the 4×4 paraxial equations $d\mathbf{Y}/dl = \mathbf{A}\cdot\mathbf{Y}$. Similar paraxial integrations are a central feature of all previously published amplitude perturbation schemes. The closest result to eq. (163) that we have found in the previous literature is that given by Coates & Chapman (1990). Their expression for $\delta\mathbf{Q}_{0L}$ does not agree with ours, because they have incorrectly perturbed the scale factor h in computing the perturbation to the Hamiltonian $\delta\mathcal{H}$; furthermore, they have not reduced the terms involving the path perturbation $\delta\mathbf{y}$, as we have done.

7.1 Realistic background models

In order to simplify our *ab initio* amplitude analysis as much as possible, we have restricted attention to a background medium that is both infinite and smooth. Despite this simplification, we are confident that the final explicit expression (163) for $\delta(\ln A) = -\frac{1}{2}|J|^{-1}\delta|J|$ is valid for an arbitrary finite, piecewise smooth background earth model with an outer free surface, an inner and outer core boundary, and other internal discontinuities. The presence of these boundaries must of course be taken into account in the computation of the traveltime Hessians $\vec{\mathbf{M}}$, $\overleftarrow{\mathbf{M}}$ and $\overleftrightarrow{\mathbf{M}}$; detailed procedures for doing this are described in Červený (1985) and banana–doughnut I. Volumetric perturbations $\delta\rho$, $\delta\sigma$ on either side of the internal or external boundaries will give rise to perturbations in the reflection and transmission coefficients that are not accounted for in eq. (163). It may also not be remiss to reiterate that we have specifically excluded the possibility of 2D perturbations in the topography of any of the boundaries.

7.2 Reciprocity redux

The ray path perturbation $\delta\mathbf{q}$, the inverse traveltime Hessian $\overleftrightarrow{\mathbf{M}}^{-1}$ and the quantity $\mathrm{tr}(\vec{\mathbf{M}}\cdot\overleftrightarrow{\mathbf{M}}^{-1}\cdot\overleftarrow{\mathbf{M}})$ are all invariant under an interchange $\mathbf{s}\leftrightarrow\mathbf{r}$ of the source and receiver:

$$\delta\mathbf{q}_{\mathrm{rs}} = \delta\mathbf{q}_{\mathrm{sr}}, \qquad \overleftrightarrow{\mathbf{M}}^{-1}_{\mathrm{rs}} = \overleftrightarrow{\mathbf{M}}^{-1}_{\mathrm{sr}}, \tag{164}$$

$$\mathrm{tr}(\vec{\mathbf{M}}\cdot\overleftrightarrow{\mathbf{M}}^{-1}\cdot\overleftarrow{\mathbf{M}})_{\mathrm{rs}} = \mathrm{tr}(\vec{\mathbf{M}}\cdot\overleftrightarrow{\mathbf{M}}^{-1}\cdot\overleftarrow{\mathbf{M}})_{\mathrm{sr}}. \tag{165}$$

It follows from eqs (164) and (165) that the ray-theoretical amplitude perturbation (163) is consistent with the principle of source–receiver reciprocity:

$$\delta(\ln A)_{\mathrm{rs}} = \delta(\ln A)_{\mathrm{sr}}. \tag{166}$$

The endpoint terms in eq. (140) are likewise invariant under a swap $\mathbf{s}\leftrightarrow\mathbf{r}$, so that the perturbed pressure response in an acoustic medium satisfies $\delta u_{\mathrm{rs}}(\omega) = \delta u_{\mathrm{sr}}(\omega)$.

7.3 Homogeneous background medium

The above ray-theoretical results can be simplified considerably in the special case of a background medium with a homogeneous slowness distribution, $\sigma =$ constant. First, all of the terms involving the cross-path gradients of the background slowness, $\boldsymbol{\nabla}_\perp\sigma$, $\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\sigma$ and $\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\sigma$, vanish. Secondly, the forward and backward propagators in such a medium can be written explicitly in the form

$$\vec{\boldsymbol{\Pi}}_{ll'} = \begin{pmatrix}\mathbf{I} & \sigma^{-1}(l-l')\mathbf{I}\\ \mathbf{0} & \mathbf{I}\end{pmatrix}, \qquad \overleftarrow{\boldsymbol{\Pi}}_{ll'} = \begin{pmatrix}\mathbf{I} & \sigma^{-1}(l'-l)\mathbf{I}\\ \mathbf{0} & \mathbf{I}\end{pmatrix}. \tag{167}$$

Upon utilizing the relations (167) we find that the initial and final slowness perturbations (136)–(137), the pointwise path perturbation (138)–(139), and the logarithmic amplitude perturbation (163) reduce to

$$\delta\mathbf{p}_{\rm s} = -\frac{1}{L}\int_0^L (L-l)\boldsymbol{\nabla}_\perp\delta\sigma\, dl, \qquad \delta\mathbf{p}_{\rm r} = \frac{1}{L}\int_0^L l\boldsymbol{\nabla}_\perp\delta\sigma\, dl, \tag{168}$$

$$\delta\mathbf{q}_l = -\frac{1}{\sigma L}\int_0^l l'(L-l)\boldsymbol{\nabla}_\perp\delta\sigma_{l'}\, dl' - \frac{1}{\sigma L}\int_l^L l(L-l')\boldsymbol{\nabla}_\perp\delta\sigma_{l'}\, dl', \tag{169}$$

$$\delta(\ln A) = \frac{1}{\sigma L}\int_0^L \delta\sigma\, dl - \frac{1}{2\sigma L}\int_0^L l(L-l)\nabla_\perp^2\delta\sigma\, dl. \tag{170}$$

Eqs (168) and (169) are reminiscent of the expressions for the takeoff and arrival angle perturbations and the path perturbation of a surface wave given by Woodhouse & Wong (1986) and Dahlen & Tromp (1998, Section 16.8.4). Note that the perturbed ray $\delta\mathbf{q}$ veers in the direction $-\boldsymbol{\nabla}_\perp\delta\sigma$, in order to find the new least-time path between the source $\mathbf{s}$ and the receiver $\mathbf{r}$. The second term in eq. (170), involving the cross-path Laplacian $\nabla_\perp^2\delta\sigma$ of the 3D slowness perturbation, has been derived by Neele *et al.* (1993); however, the first term, which is simply the path average of the fractional slowness perturbation $\delta\sigma/\sigma$, has been ignored in their analysis. It is evident from consideration of the simple case of a uniform perturbation that this term must be present. In the case of $\delta\sigma =$ constant, the path average in eq. (170) is cancelled by the slowness endpoint contributions in eq. (140), so that the overall amplitude perturbation reduces to $\delta(\ln A)_{\rm total} = \frac{1}{2}(\delta\rho_{\rm s}/\rho_{\rm s} + \delta\rho_{\rm r}/\rho_{\rm r})$, in accordance with an elementary analysis.

The dependence of the ray-theoretical amplitude perturbation $\delta(\ln A)$ upon the cross-path slowness curvature $\nabla_\perp^2\delta\sigma$ is consistent with the expected character of ray-tube area variations: propagation through a slow lens or channel, $\nabla_\perp^2\delta\sigma < 0$, gives rise to geometrical focusing and pulse amplification, $\delta(\ln A) > 0$, whereas propagation through a fast lens or channel, $\nabla_\perp^2\delta\sigma > 0$, gives rise to defocusing and deamplification, $\delta(\ln A) < 0$. The presence of the quadratic along-ray weighting term $l(L-l)$ in eq. (170) indicates that curvature variations near the source $\mathbf{s}$ and receiver $\mathbf{r}$ exert a negligible influence, whereas those near the midpoint of the ray path are most significant. Similar remarks apply to the cross-path curvature term $\overleftrightarrow{\mathbf{M}}^{-1} : \boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\delta\sigma$ in the general expression (163) for $\delta(\ln A)$, as a consequence of the endpoint limiting relations (84). Whether positive or negative slowness curvature $\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\delta\sigma$ in a given cross-path direction gives rise to focusing or defocusing depends in the general case upon the signature of the Hessian, $\operatorname{sig}\overleftrightarrow{\mathbf{M}} = 2, 0, -2$.

7.4 Good enough for government work?

We turn at last to a discussion of the question posed in the introduction—*does the 3D amplitude kernel formulation reduce to the corresponding result obtained using geometrical ray perturbation theory in the limit of infinite frequency*, as expected? Upon comparing the ray-theoretical result eq. (163), which we rewrite in an annotated form below,

$$\begin{aligned}\delta(\ln A) = {} & \underbrace{\frac{1}{2}\int_0^L \sigma^{-2}\delta\sigma\,\operatorname{tr}(\vec{\mathbf{M}}\cdot\overleftrightarrow{\mathbf{M}}^{-1}\cdot\overleftarrow{\mathbf{M}})\, dl}_{\text{oops}} \underbrace{-\frac{1}{2}\int_0^L \overleftrightarrow{\mathbf{M}}^{-1} : \Big(\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\delta\sigma - 2\sigma^{-1}\boldsymbol{\nabla}_\perp\sigma\boldsymbol{\nabla}_\perp\delta\sigma\Big)\, dl}_{\text{eq.(53)}} \\ & \underbrace{-\frac{1}{2}\int_0^L \overleftrightarrow{\mathbf{M}}^{-1} : \Big[\delta\mathbf{q}\cdot\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\sigma - 2\sigma^{-1}(\delta\mathbf{q}\cdot\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\sigma)\boldsymbol{\nabla}_\perp\sigma - \sigma^{-1}(\delta\mathbf{q}\cdot\boldsymbol{\nabla}_\perp\sigma)\Big(\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\sigma + 3\sigma^{-1}\boldsymbol{\nabla}_\perp\sigma\boldsymbol{\nabla}_\perp\sigma\Big)\Big]\, dl}_{\text{oops}},\end{aligned} \tag{171}$$

with eq. (53) we see that the answer is ... almost, but not quite. The second integral involving the second and first cross-path gradients $\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\delta\sigma$ and $\boldsymbol{\nabla}_\perp\delta\sigma$ is identical; however, the first 'oops' integral involving the undifferentiated slowness perturbation $\delta\sigma$ and the final 'oops' integral involving the path perturbation $\delta\mathbf{q}$ are unaccounted for in the $\omega\to\infty$ limit of the 3D Born kernel formulation. Since eq. (138) stipulates that the path perturbation $\delta\mathbf{q}$ depends upon the first cross-path gradient $\boldsymbol{\nabla}_\perp\delta\sigma$, we are obliged to admit that eqs (53) and (171) agree to order $\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\delta\sigma$ but not to order $\boldsymbol{\nabla}_\perp\delta\sigma$ or $\delta\sigma$. The latter terms are of the same order as the acoustic endpoint terms in eq. (140), which are also unaccounted for in eq. (171). In a more general elastic earth model, the endpoint terms will depend upon the slowness perturbations $\delta\mathbf{p}_{\rm s}$ and $\delta\mathbf{p}_{\rm r}$, which are of order $\boldsymbol{\nabla}_\perp\delta\sigma$, as seen in eqs (136) and (137).

As already noted, we expect on the basis of scaling considerations that the curvature term $\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\delta\sigma$ will dominate the focusing and defocusing contribution to $\delta(\ln A)$ whenever the background slowness σ is much smoother than the 3D perturbation $\delta\sigma$. In this case, at least, it is legitimate to describe the paraxial, forward-scattering amplitude Fréchet kernel formulation in eqs (5) and (21) as the natural finite-frequency extension of linearized geometrical ray theory, eq. (163).

8 CONCLUSION

In summary, we have developed a 3D Fréchet kernel K_A for rms body-wave amplitude measurements, as a complement to the 'banana–doughnut' Fréchet kernel K_T for cross-correlation traveltimes. The two kernels express the sensitivity of logarithmic amplitude and traveltime

perturbations $\delta(\ln A)$ and δT to 3D seismic slowness variations $\delta\sigma/\sigma$. Roughly speaking, the amplitude kernel K_A exhibits a local maximum at points $0 \le l \le L$ along the central source-to-receiver ray where the traveltime kernel K_T is identically zero, and vice versa. In the asymptotic limit of infinite frequency, $\omega \to \infty$, both the amplitude and traveltime Fréchet kernel formulations are consistent with geometrical ray perturbation theory. The dominant ray-theoretical dependence of $\delta(\ln A)$ upon the cross-path curvature of the slowness perturbation $\nabla_\perp \nabla_\perp \delta\sigma$ is captured by the $\omega \to \infty$ asymptotic reduction of the 3D amplitude integral, but not the weaker dependence upon a welter of $\nabla_\perp \delta\sigma$ and $\delta\sigma$ terms.

ACKNOWLEDGMENTS

We wish to express our profound gratitude to one of the reviewers of this paper, Veronique Farra, for pointing out a significant error in our initial analysis of the ray-theoretical amplitude perturbation. Her diligence, insight and deep knowledge of seismic ray theory are greatly appreciated. We also thank Roel Snieder, who provided a second very constructive review, and Shu-Huei Hung, who wrote the software used to compute the 3D Fréchet kernels plotted in Figs 4–9. Financial support for this work has been provided by the US National Science Foundation under Grants EAR-9725496 and EAR-0105387.

REFERENCES

Aki, K. & Richards, P.G., 1980. *Quantitative Seismology: Theory and Methods,* Freeman, San Francisco.

Červený, V., 1985. The application of ray tracing to the numerical modeling of seismic wavefields in complex structures, in, *Seismic Shear Waves, Part A: Theory, Handbook of Geophysical Exploration,* Vol. 15A, pp. 1–124, ed. Dohr, G.P., Geophysical Press, London.

Červený, V. & Hron, F., 1980. The ray series method and dynamic ray tracing systems for 3D inhomogeneous media, *Bull. seism. Soc. Am.,* **70**, 47–77.

Coates, R.T. & Chapman, C.H., 1990. Ray perturbation theory and the Born approximation, *Geophys. J. Int.,* **100**, 379–392.

Dahlen, F.A. & Tromp, J., 1998. *Theoretical Global Seismology,* Princeton University Press, Princeton.

Dahlen, F.A., Hung, S.-H. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes—I. Theory, *Geophys. J. Int.,* **141**, 157–174.

Dziewonski, A.M. & Anderson, D.L., 1981. Preliminary reference Earth model, *Phys. Earth planet. Inter.,* **25**, 297–356.

Farra, V. & Madariaga, R., 1987. Sesimic waveform modeling in heterogeneous media by ray perturbation theory, *J. geophys. Res.,* **92**, 2697–2712.

Gilbert, F. & Backus, G.E., 1966. Propagator matrices in elastic wave and vibration problems, *Geophysics,* **31**, 326–333.

Goldstein, H., 1980. *Classical Mechanics,* Addison-Wesley, Reading, MA.

Horn, R.A. & Johnson, C.A., 1985. *Matrix Analysis,* Cambridge University Press, Cambridge.

Hung, S.-H., Dahlen, F.A. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes—II. Examples, *Geophys. J. Int.,* **141**, 175–203.

Kuo, B.-Y., Forsyth, D.W. & Wysession, M., 1987. Lateral heterogeneity and azimuthal anisotropy in the North Atlantic determined from *SS–S* differential traveltimes, *J. geophys. Res.,* **92**, 6421–6436.

Lanczos, C., 1962. *The Variational Principles of Mechanics,* University of Toronto Press, Toronto.

Liu, X.-F. & Tromp, J., 1996. Uniformly valid body-wave theory, *Geophys. J. Int.,* **127**, 461–491.

Masters, G., Johnson, S., Laske, G. & Bolton, H., 1996. A shear-velocity model of the mantle, *Phil. Trans. R. Soc. Lond.,* A., **354**, 1385–1411.

Morse, P.M. & Ingard, K.U., 1968. *Theoretical Acoustics,* McGraw-Hill, New York.

Neele, F., VanDecar, J.C. & Snieder, R., 1993. A formalism for including amplitude data in tomographic inversions, *Geophys. J. Int.,* **115**, 482–496.

Nowack, R.L. & Lutter, W.J., 1988. Linearized rays, amplitude and inversion, *Pageoph,* **128**, 401–421.

Sheehan, A. & Solomon, S., 1991. Joint inversion of shear wave traveltime residuals and geoid and depth anomalies for long-wavelength variations in upper mantle temperature and composition along the Mid-Atlantic Ridge, *J. geophys. Res.,* **96**, 19 981–20 009.

Snieder, R., 1987. On the connection between ray theory and scattering theory for surface waves, in, *Mathematical Geophysics,* pp. 77–84, eds Vlaar, N.J., Nolet, G., Wortel, M.J.R. & Cloetingh, S.A.P.L., Reidel, Dordrecht.

Snieder, R. & Chapman, C.H., 1998. The reciprocity properties of geometrical spreading, *Geophys. J. Int.,* **132**, 89–95.

Snieder, R. & Lomax, A., 1996. Wavefield smoothing and the effect of rough velocity perturbations on arrival times and amplitudes, *Geophys. J. Int.,* **125**, 796–812.

Snieder, R. & Sambridge, M., 1992. Ray perturbation theory for traveltimes and ray paths in 3D heterogeneous media, *Geophys. J. Int.,* **109**, 294–322.

Thomson, C., 1983. Ray-theoretical amplitude inversion for laterally varying structure below NORSAR, *Geophys. J. R. astr. Soc.,* **74**, 525–558.

Whitham, G.B., 1974. *Linear and Nonlinear Waves,* Wiley, New York.

Woodhouse, J.H. & Wong, Y.K., 1986. Amplitude, phase and path anomalies of mantle waves, *Geophys. J. R. astr. Soc.,* **87**, 753–773.

Woodward, R.L. & Masters, G., 1991. Global upper mantle structure from long-period differential traveltimes, *J. geophys. Res.,* **96**, 6351–6377.