Wavefield reciprocity and local optimization in remote sensing

MAARTEN V. DE HOOP *and* ADRIANUS T. DE HOOP *
Center for Wave Phenomena,
Colorado School of Mines,
Golden CO 80401, USA.

April 3, 1997

Abstract

A general local optimization approach to imaging/inversion in remote sensing is presented. The approach is based on the time-domain reciprocity theorems of the time-convolution and the time-correlation types that apply to wavefields in linear, time-invariant configurations. The wavefields satisfy a system of linear, first-order, partial differential equations in space-time of a general class to which acoustic waves in fluids, elastic waves in solids and electromagnetic waves belong. First, the reciprocity theorems are derived. Next, the remote sensing problem is formulated as a reconstruction problem of the contrast-in-medium parameters in a subdomain of the configuration space where the medium properties differ by a certain amount from the ones in a known embedding. Finally, an optimization procedure is developed that leads to a guaranteed decrease in the mismatch of the measured field in a chosen domain of observation. The scheme is shown to encompass – and partially unify – the various known imaging/inversion procedures that are in practical use in seismic and electromagnetic (radar) geophysical exploration and environmental monitoring.

*On leave from Laboratory of Electromagnetic Research, Faculty of Electrical Engineering, Delft University of Technology, the Netherlands.

1 Introduction

In this paper, we discuss the wave-theoretical framework of local optimization approaches to the inverse scattering problem in remote sensing. In the discussion, we cover the issues of matching observed data in the least squares (weak) sense with simulations, of adjoint scattering operators and states, of preconditioning and of differential semblance. The fundamental reciprocity theorems of the time-convolution and time-correlation types [1] form the basis of our analysis; medium perturbations are introduced through the contrast-source formulation; a local Born approximation is employed to generate a series expansion for the actual reconstruction.

An important aspect of taking the reciprocity relation of the time-correlation type as a point of departure is that it induces a proper norm in the function space of the pertaining wavefields and thus leads, via the iterative minimization according to a chosen criterion, to a 'minimum-norm' gradient technique with a guaranteed decreasing error in the data fit (Lions [2]). Special attention is paid to how the different wave types contribute. Within the framework of reciprocity, there is still some freedom in the choice of norm. This choice allows for a statistical reformulation of the inverse problem.

The concept of data fitting with a view to resolution analysis has been introduced by Backus and Gilbert [3]. It has been further exploited and developed into an optimization procedure in a large number of papers, including the ones by Bamberger *et al.* [4], Tarantola and Valette [5], Lailly [6], Kolb *et al.* [7], Kennett and Williamson [8], Nolet [9] and Symes [10]. Tarantola's [11] observation that inversion can be expressed in terms of more conventional image processing methods is a recurring motif in the literature that applies to almost all common approaches to the inverse problem. The fundamental aspects of the optimization method are discussed here and the results are integrated with the imaging conditions of migration/inversion, both the ones based on the Generalized Radon Transform (see Miller *et al.* [12] and Esmersoy and Oristaglio [13]), and the ones based on full-wave theory (see Claerbout [14, 15], Berkhout [16], and Wapenaar and Berkhout [17]), by means of preconditioning. Sophisticated preconditioning can be designed with the aid of linearized inversion. For early attempts to transform seismic imaging into linearized inversion, we refer to Stolt and Weglein [18] and Bleistein [19]. For a general overview of weak inverse scattering theory, we refer to Parker [20].

The approach to inverse scattering discussed in this paper is based on full wave theory and thus requires a considerable computational effort. However, asymptotic and other approximate techniques can be employed to enhance the efficiency of the methods discussed here, in particular during the first iteration.

2 The basic field equations

We consider linear acoustic, elastic or electromagnetic wave motion in some subdomain D of three-dimensional Euclidean space $\mathbb{R}^3$ [21, 22]. The configurations in which the wave motion is considered to be present, are assumed to be time invariant. The wave quantities involved are found to satisfy certain reciprocity properties which will be taken as point of departure for our further considerations. Now, for the indicated type of configuration, there prove to be two kinds of reciprocity theorem: one of the time-convolution type, the other of the time-correlation type. Several operations on the wave quantities will occur throughout the paper. First, we shall introduce their notation.

Notation

Cartesian coordinates $\boldsymbol{x} = \{x_1, x_2, x_3\}$ are used to specify position, with respect to the base $\{\boldsymbol{i}_1, \boldsymbol{i}_2, \boldsymbol{i}_3\}$; t is the time coordinate. Differentiation with respect to x_p is denoted by ∂_p and the gradient vector by $\nabla_{\boldsymbol{x}}$; ∂_t is a reserved symbol for differentiation with respect to t. The subscript notation for the vectorial and tensorial quantities occurring in the wave motion will be used whenever appropriate; the lowercase subscripts are to be assigned the values 1, 2 and 3.

The characteristic function of the domain D is denoted by χ_D and is given by

$$\chi_D(\boldsymbol{x}) = \{1, \tfrac{1}{2}, 0\} \quad \text{for} \quad \boldsymbol{x} \in \{D, \partial D, D'\} \; , \tag{2.1}$$

where ∂D is the boundary of D and D' is the complement of $D \cup \partial D$ in $\mathbb{R}^3$.

Let $F = F(\boldsymbol{x}, t)$ denote any space-time function. Then, the *time reversal* operator is defined by

$$\mathsf{T}(F)(\boldsymbol{x}, t) = F(\boldsymbol{x}, -t) \; . \tag{2.2}$$

It has the property

$$\partial_t \mathsf{T}(F) = -\mathsf{T}(\partial_t F) \; . \tag{2.3}$$

Let $Q(\boldsymbol{x}, t)$ denote another space-time function, then the *time convolution* of F and Q is denoted by

$$\mathsf{C}_t(F, Q)(\boldsymbol{x}, t) = \int_{t' \in \mathbb{R}} F(\boldsymbol{x}, t') Q(\boldsymbol{x}, t - t') \; \mathrm{d}t' \; . \tag{2.4}$$

It has the properties

$$\mathsf{C}_t(F,Q) = \mathsf{C}_t(Q,F) \;, \tag{2.5}$$

$$\mathsf{C}_t(\mathsf{T}(F),\mathsf{T}(Q)) = \mathsf{T}\mathsf{C}_t(F,Q) \;, \tag{2.6}$$

$$\partial_t \mathsf{C}_t(F,Q) = \mathsf{C}_t(F,\partial_t Q) = \mathsf{C}_t(\partial_t F,Q) \;. \tag{2.7}$$

A compound time convolution is denoted by a single operator C_t acting on the constituting functions; the order in which the latter appear is immaterial. For example,

$$\mathsf{C}_t(F,\mathsf{C}_t(G,Q)) = \mathsf{C}_t(F,G,Q) \;.$$

The *time deconvolution* is denoted as C_t^{-1}. For example,

$$\mathsf{C}_t^{-1}(F,Q) = \mathsf{C}_t(F,Q') \quad \text{with} \quad \mathsf{C}_t(Q',Q)(t) = \delta_0(t) \;.$$

The *time correlation* of F and Q is denoted by

$$\mathsf{R}_t(F,Q)(\boldsymbol{x},t) = \int_{t'\in\mathbb{R}} F(\boldsymbol{x},t')Q(\boldsymbol{x},t'-t)\;\mathrm{d}t' \;. \tag{2.8}$$

It has the properties

$$\mathsf{R}_t(F,Q) = \mathsf{C}_t(F,\mathsf{T}(Q)) \;, \tag{2.9}$$

$$\mathsf{R}_t(Q,F) = \mathsf{T}\mathsf{R}_t(F,Q) \;, \tag{2.10}$$

$$\partial_t \mathsf{R}_t(F,Q) = -\mathsf{R}_t(F,\partial_t Q) = \mathsf{R}_t(\partial_t F,Q) \;. \tag{2.11}$$

Since in a compound time correlation the order of the constituting functions is of importance, they are profitably rewritten as convolutions by using Eq.(2.9).

Throughout the paper, we will employ a notation close to the one developed by Woodhouse [23], and refined by De Hoop [24] and De Hoop and De Hoop [25].

The hyperbolic system of equations

Let the *field matrix* $F_P = F_P(\boldsymbol{x},t)$ of the wave motion be composed of the components of the two wavefield quantities whose inner product represents the area density of power flow (Poynting vector). Then, F_P satisfies a system of linear, first-order, partial differential equations of the general form

$$(\mathcal{D}_{IP} + M_{IP}\partial_t)\,F_P = Q_I \;, \quad \mathcal{D}_{IP} = \mathcal{D}_{IP}(\nabla) \;,\; M_{IP} = M_{IP}(\boldsymbol{x}) \;, \tag{2.12}$$

where uppercase Latin subscripts are used to denote the pertaining matrix elements and the summation convention for repeated subscripts applies. In Equation (2.12), $\mathcal{D}_{IP}$ is a symmetric, block off-diagonal *spatial differentiation operator matrix* that contains the operator ∂_p in a homogeneous linear fashion that is specific for each type of wave motion under consideration, M_{IP} is the *medium matrix* that is representative for the properties of the (arbitrarily inhomogeneous, anisotropic) media in which the waves propagate and $Q_I = Q_I(\boldsymbol{x}, t)$ is the *volume source density matrix* that is representative for the action of the volume sources that generate the wavefield. Also, surface sources will be included in the disussion. The corresponding *surface source density matrix* is denoted by

$$q_I = [\mathcal{N}_{IP} F_P]_-^+ = \mathcal{N}_{IP} F_P|^+ + \mathcal{N}_{IP} F_P|^- \ , \tag{2.13}$$

where $[\cdot]_-^+$ denotes the jump across the support of the surface source distributions and $\mathcal{N}_{IP}$ is the *unit normal operator* that arises from replacing ∂_p in $\mathcal{D}_{IP}$ by n_p, where $\boldsymbol{n}$ is, on each of the two faces of the surface, the unit vector along the normal oriented away from the domain that surrounds that surface, $\mathcal{N}_{IP} = \mathcal{D}_{IP}(\boldsymbol{n})$.

The medium parameters are assumed to be piecewise continuous. Across a surface of discontinuity in medium properties, the parameters may jump by finite amounts. On the assumption that the interface is passive (i.e., free from surface sources) and that the wavefield quantities must remain bounded on either side of the interface, the wavefield must satisfy the boundary condition of the continuity type

$$\mathcal{N}_{IP} F_P \quad \text{is continuous across source-free interface} \ . \tag{2.14}$$

We have

- for acoustic waves in fluids, $F_P = [p, v_1, v_2, v_3]^T$, where p = acoustic pressure and v_r = particle velocity, and $Q_I = [q, f_1, f_2, f_3]^T$, where q = volume source density of injection rate and f_k = volume source density of force;

- for elastic waves in solids, $F_P = [v_1, v_2, v_3, -\tau_{11}, -\tau_{12}, -\tau_{13}, -\tau_{21}, -\tau_{22}, -\tau_{23}, -\tau_{31}, -\tau_{32}, -\tau_{33}]^T$, where v_r = particle velocity and τ_{pq} = dynamic stress, and $Q_I = [f_1, f_2, f_3, h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32}, h_{33}]^T$, where f_k = volume source density of force and h_{ij} = volume source density of deformation rate;

- for electromagnetic fields, $F_P = [E_1, E_2, E_3, 0, -H_3, H_2, H_3, 0, -H_1, -H_2, H_1, 0]^T$, where E_r = electric field strength and H_p = magnetic field strength, and $Q_I = [-J_1, -J_2, -J_3, 0, \frac{1}{2}K_3, -\frac{1}{2}K_2, -\frac{1}{2}K_3, 0, \frac{1}{2}K_1, \frac{1}{2}K_2, -\frac{1}{2}K_1, 0]^T$, where J_k = volume source density of electric current and K_j = volume source density of magnetic current.

The structure of $\mathcal{D}_{IP}$ for the three types of wave fields is given in Appendix A; the medium matrices are discussed in Appendix B.

The reciprocity concatenation matrices

In the reciprocity theorems to be discussed in Section 3, two diagonal matrices δ^-_{QI} and δ^+_{QI} occur that concatenate out of the wavefields pertaining to two admissible states their interaction. For acoustic waves in fluids the diagonal matrix δ^-_{QI} is given by $\mathrm{diag}[1,-1,-1,-1]$, for elastic waves in solids by $\mathrm{diag}[1,1,1,-1,-1,-1,-1,-1,-1,-1,-1,-1]$ and for electromagnetic waves also by $\mathrm{diag}[1,1,1,-1,-1,-1,-1,-1,-1,-1,-1,-1]$. The diagonal matrix δ^+_{QI} is just the unit matrix: $\delta^+_{QI} = 1$ for $Q = I$, $\delta^+_{QI} = 0$ for $Q \neq I$.

For the reciprocity theorem of the *time-convolution* type to hold, a necessary and sufficient condition proves to be

$$\delta^-_{QI}\mathcal{D}_{IP} = -\delta^-_{PJ}\mathcal{D}_{JQ} \ . \tag{2.15}$$

This condition requires that the block-diagonal part of $\mathcal{D}_{IP}$ is anti-symmetric and that its block off-diagonal part is symmetric.

For the reciprocity theorem of the *time-correlation* type to hold, a necessary and sufficient condition proves to be

$$\delta^+_{QI}\mathcal{D}_{IP} = \delta^+_{PJ}\mathcal{D}_{JQ} \ . \tag{2.16}$$

This condition requires that $\mathcal{D}_{IP}$ is symmetric. The two conditions are independent, but if they are satisfied simultaneously, $\mathcal{D}_{IP}$ is a symmetric, block off-diagonal matrix operator. For the three types of wave motion considered in this paper, this is indeed the case. It is therefore conjectured that the indicated structure of the spatial differential matrix operator could be fundamental to a system of first-order partial differential equations in general to be representative for a wave motion. It is noted that the medium matrix M_{IP} is not subjected to any restriction of this kind.

Asymptotic ray theory: eikonal and transport equations

To understand the singular behavior of the wave solution of system (2.12) we employ the standard asymptotic expansion for the field matrix,

$$F_P(\boldsymbol{x},t) = \sum_{\text{modes,paths}} \sum_{n=0}^{\infty} F_P^{(n)}(\boldsymbol{x})\, f_n(t - \tau(\boldsymbol{x}))\ , \tag{2.17}$$

where τ is the traveltime along a characteristic, assuming there is single source signature the derivative of which coincides with f_0 and that generates the field, and

$$(f_n)' = f_{n-1} \ , \quad n = 1, 2, \cdots \ . \tag{2.18}$$

For notational convenience, set $F_P^{(n)} \equiv 0$ for $n < 0$. Substituting the asymptotic expansion into Eq.(2.12) away from any source distributions yields

$$\mathcal{D}_{IP}(\nabla) \, F_P^{(n-1)} + [\mathcal{D}_{IP}(\boldsymbol{\gamma}) + M_{IP}(\boldsymbol{x})] \, F_P^{(n)} = 0 \ , \quad n = 0, 1, \cdots \ , \tag{2.19}$$

where

$$\boldsymbol{\gamma} = \nabla \tau \tag{2.20}$$

is the so-called *slowness vector*. Collecting the zero-order terms, yields the Christoffel equation

$$[\mathcal{D}_{IP}(\boldsymbol{\gamma}) + M_{IP}(\boldsymbol{x})] \, F_P^{(0)} = 0 \ . \tag{2.21}$$

The zero-order field constituent can be written in the form

$$F_P^{(0)} = A^{(0)} \, \Xi_P \ , \tag{2.22}$$

where $A^{(0)}$ is a scalar amplitude. Substituting Eq.(2.22) into Eq.(2.21) leads to the eigenvalue equation

$$[\mathcal{D}_{IP}(\boldsymbol{\gamma}) + M_{IP}(\boldsymbol{x})] \, \Xi_P = 0 \ , \quad \Xi_P = \Xi_P(\boldsymbol{x}, \boldsymbol{\gamma}) \ . \tag{2.23}$$

Non-trivial solutions to Eq.(2.23) must satisfy

$$\det[\mathcal{D}_{IP}(\boldsymbol{\gamma}) + M_{IP}(\boldsymbol{x})] = 0 \ , \tag{2.24}$$

which equation is equivalent to the *eikonal* equation. Collecting the first-order terms from Eq.(2.19) and contracting the result with $F_I^{(0)}$, leads in view of Eq.(2.21) to

$$F_I^{(0)}(\mathcal{D}_{IP}(\nabla) \, F_P^{(0)}) = 0 \ . \tag{2.25}$$

Using the symmetry of $\mathcal{D}$ (Eq.(2.16)) and imposing that the medium matrix is symmetric as well, this equation can be rewritten as a conservation law,

$$\tfrac{1}{2}\partial_p[F_I^{(0)} \mathcal{D}_{IP}(\boldsymbol{i}_p) \, F_P^{(0)}] = 0 \ . \tag{2.26}$$

In here, we identify the vectorial quantity,

$$\mathrm{P}_p = \tfrac{1}{2} F_Q \mathcal{D}_{QP}(\boldsymbol{i}_p) \, F_P \ , \tag{2.27}$$

which, in the energy balance to be discussed later, appears to be the Poynting vector. Equation (2.26) is equivalent to the *transport* equation, see Chapman and Coates [26].

Point-source solutions, Green's tensors

In view of the linearity of the wave motion, the principle of superposition ensures that the wavefield F_P that is generated by the volume source distribution Q_I and the surface source distributions q_I can be written as the superposition of point-source contributions through the use of a Green's tensor. The latter is a solution of the system of differential equations

$$(\mathcal{D}_{IP} + M_{IP}\partial_t)\, G_{PI'} = \delta^{+}_{II'}\delta_{\boldsymbol{x}'}\delta_{t'} \;, \tag{2.28}$$

where $\delta_{II'}$ is the unit matrix and $\delta.(.)$ is the Dirac distribution. In view of the time invariance of the medium, the Green's tensor depends on t and t' only through the difference $t - t'$, i.e.,

$$G_{PI'} = G_{PI'}(\boldsymbol{x}, \boldsymbol{x}', t, t') = G_{PI'}(\boldsymbol{x}, \boldsymbol{x}', t - t') \;.$$

The Green's tensor plays an important rôle in the reconstruction problem. If it is used to represent actual physical (i.e., causal) wavefields, the corresponding Green's tensor is to be taken as the causal solution of Equation (2.28).

3 The reciprocity theorems

In the wavefield reciprocity theorems certain *interaction quantities* are considered that are representative for the interaction between two admissible states of the pertaining wavefield in a given (proper or improper) subdomain D of $\mathbb{R}^3$. Each of the two states has its own medium and its own volume source distribution. Let the superscripts Y and Z indicate the two states, then the wavefields in the two states are related to their respective sources via

$$(\mathcal{D}_{IP} + M^{Y}_{IP}\partial_t)\, F^{Y}_{P} \;=\; Q^{Y}_{I} \;, \tag{3.1}$$

$$(\mathcal{D}_{JQ} + M^{Z}_{JQ}\partial_t)\, F^{Z}_{Q} \;=\; Q^{Z}_{J} \;. \tag{3.2}$$

and

$$[\mathcal{N}_{IP} F^{Y}_{P}]^{+}_{-} \;=\; q^{Y}_{I} \;, \tag{3.3}$$

$$[\mathcal{N}_{JQ} F^{Z}_{Q}]^{+}_{-} \;=\; q^{Z}_{J} \;. \tag{3.4}$$

Further, for each of the two states the boundary condition of the continuity type

$$\mathcal{N}_{IP} F^{Y}_{P} \quad \text{is continuous across source-free interface} \;, \tag{3.5}$$

$$\mathcal{N}_{JQ} F^{Z}_{Q} \quad \text{is continuous across source-free interface} \tag{3.6}$$

holds.

The reciprocity theorem of the time-convolution type

The local interaction quantity to be considered in the reciprocity theorem of the time-convolution type is

$$\delta^-_{QI}\,\mathsf{C}_t(\mathcal{D}_{IP}F^Y_P, F^Z_Q) - \delta^-_{PJ}\,\mathsf{C}_t(F^Y_P, \mathcal{D}_{JQ}F^Z_Q) = \delta^-_{QI}\mathcal{D}_{IP}\,\mathsf{C}_t(F^Y_P, F^Z_Q)\ ,$$

where the property of Eq.(2.15) has been used. With the aid of Eqs.(2.7), (3.1) and (3.2) this expression is rewritten as

$$\delta^-_{QI}\mathcal{D}_{IP}\,\mathsf{C}_t(F^Y_P, F^Z_Q) + (\delta^-_{QI}M^Y_{IP} - \delta^-_{PJ}M^Z_{JQ})\,\partial_t\mathsf{C}_t(F^Y_P, F^Z_Q) = \delta^-_{QI}\,\mathsf{C}_t(Q^Y_I, F^Z_Q) - \delta^-_{PJ}\,\mathsf{C}_t(F^Y_P, Q^Z_J)\ . \tag{3.7}$$

Equation (3.7) is the local form of the reciprocity theorem of the time-convolution type. The global form, for the domain D, of this theorem follows upon integrating Eq.(3.7) over the domain D and applying Gauss' divergence theorem to the first term on the left-hand side over each subdomain of D where the field quantities are continuously differentiable. Adding the contributions from these subdomains, the contributions from source-free interfaces of discontinuity in medium properties in the interior of D cancel in view of the boundary conditions given in Eqs.(3.5) and (3.6) and only surface integrals over the support S of the surface sources and the boundary ∂D of D remain. The result is

$$\begin{aligned}\int_D \left[\delta^-_{QI}\,\mathsf{C}_t(Q^Y_I, F^Z_Q) - \delta^-_{PJ}\,\mathsf{C}_t(F^Y_P, Q^Z_J)\right]\,\mathrm{d}V(\boldsymbol{x}) &= \int_{\partial D}\delta^-_{QI}\mathcal{N}_{IP}\,\mathsf{C}_t(F^Y_P, F^Z_Q)\,\mathrm{d}A(\boldsymbol{x}) \\ &+ \int_S \delta^-_{QI}\,[\mathcal{N}_{IP}\,\mathsf{C}_t(F^Y_P, F^Z_Q)]^+_-\,\mathrm{d}A(\boldsymbol{x}) \\ &+ \int_D (\delta^-_{QI}M^Y_{IP} - \delta^-_{PJ}M^Z_{JQ})\,\partial_t\mathsf{C}_t(F^Y_P, F^Z_Q)\,\mathrm{d}V(\boldsymbol{x})\ .\end{aligned} \tag{3.8}$$

Equation (3.8) is the global form, for the domain D, of the reciprocity theorem of the time-convolution type. The terms in this equation define bi-linear forms over surfaces and volumes.

The terms in Eqs.(3.7) and (3.8) containing the medium matrices define the contrast-in-medium contributions to the interaction of the two states. They vanish at those positions where $\delta^-_{QI}M^Y_{IP} - \delta^-_{PJ}M^Z_{JQ} = 0$, i.e., where $M^Y_{QP} = M^Z_{PQ}$. If this condition holds, the media in the two states are denoted as each other's *adjoints*. If the condition holds for one and the same medium, such a medium is denoted as *self-adjoint*. An isotropic medium is always self-adjoint. The terms containing the volume or surface source densities yield the contribution from the volume or surface sources, respectively, to the interaction of the two states. They vanish at source-free positions.

In a number of applications Eq.(3.8) will be applied to the entire $\mathbb{R}^3$. Then, outside some sphere $\mathcal{S}(\mathcal{O},\Delta_0)$ with radius Δ_0 and center at the origin $\mathcal{O}$ of the chosen reference frame, the media in the two states will be assumed to be the same and homogenous as well as isotropic. For such a medium, the tensor Green's function is known analytically and in particular the causal and anti-causal source-type integral representations are known explicitly. For the application of Eq.(3.8) to the entire $\mathbb{R}^3$, the theorem will be first applied to a sphere $\mathcal{S}(\mathcal{O},\Delta)$ of radius Δ and center at the origin $\mathcal{O}$ of the chosen reference frame and the limit $\Delta \to \infty$ will be taken. If, now, in both states the wavefields are causally related to the action of their volume or surface source distributions (assumed to have bounded supports), the integral over $\mathcal{S}(\mathcal{O},\Delta)$ vanishes as $\Delta \to \infty$. However, if one of the two states is causally related to the action of its volume or surface sources and the other anti-causally, the integral over $\mathcal{S}(\mathcal{O},\Delta)$ does not vanish as $\Delta \to \infty$, but has a constant value for sufficiently large values of Δ.

The reciprocity theorem of the time-correlation type

The local interaction quantity to be considered in the reciprocity theorem of the time-correlation type is

$$\delta^{+}_{QI}\,\mathsf{R}_t(\mathcal{D}_{IP}F^Y_P, F^Z_Q) + \delta^{+}_{PJ}\,\mathsf{R}_t(F^Y_P, \mathcal{D}_{JQ}F^Z_Q) = \delta^{+}_{QI}\mathcal{D}_{IP}\,\mathsf{R}_t(F^Y_P, F^Z_Q)\ ,$$

where the property of Eq.(2.16) has been used. With the aid of Eqs.(2.11), (3.1) and (3.2) this expression is rewritten as

$$\begin{aligned}\delta^{+}_{QI}\mathcal{D}_{IP}\,\mathsf{R}_t(F^Y_P, F^Z_Q) + (\delta^{+}_{QI}M^Y_{IP} - \delta^{+}_{PJ}M^Z_{JQ})\,\partial_t\mathsf{R}_t(F^Y_P, F^Z_Q)\\ = \delta^{+}_{QI}\,\mathsf{R}_t(Q^Y_I, F^Z_Q) + \delta^{+}_{PJ}\,\mathsf{R}_t(F^Y_P, Q^Z_J)\ .\end{aligned} \tag{3.9}$$

Equation (3.9) is the local form of the reciprocity theorem of the time-correlation type. The global form, for the domain D, of this theorem follows upon integrating Eq.(3.9) over the domain D and applying Gauss' divergence theorem to the first term on the left-hand side over each subdomain of D where the field quantities are continuously differentiable. Adding the contributions from these subdomains, the contributions from interfaces of discontinuity in medium properties in the interior of D cancel in view of the boundary conditions given in Eqs.(3.5) and (3.6) and only surface integrals over the support S of the surface sources and the boundary ∂D

of D remain. The result is

$$\int_D \left[\delta^+_{QI}\,\mathsf{R}_t(Q^Y_I, F^Z_Q) + \delta^+_{PJ}\,\mathsf{R}_t(F^Y_P, Q^Z_J)\right]\,\mathrm{d}V(\boldsymbol{x})$$

$$= \int_{\partial D} \delta^+_{QI}\mathcal{N}_{IP}\,\mathsf{R}_t(F^Y_P, F^Z_Q)\,\mathrm{d}A(\boldsymbol{x})$$

$$+ \int_S \delta^+_{QI}\,[\mathcal{N}_{IP}\,\mathsf{R}_t(F^Y_P, F^Z_Q)]^+_-\,\mathrm{d}A(\boldsymbol{x})$$

$$+ \int_D (\delta^+_{QI}M^Y_{IP} - \delta^+_{PJ}M^Z_{JQ})\,\partial_t \mathsf{R}_t(F^Y_P, F^Z_Q)\,\mathrm{d}V(\boldsymbol{x})\ . \tag{3.10}$$

Equation (3.10) is the global form, for the domain D, of the reciprocity theorem of the time-correlation type. It is observed that the terms occurring in this equation define a bilinear form, which induces a norm and a metric in $L^2(\mathbb{R}^4)$ for zero correlation time shift; $\mathsf{R}_t(.,.)(\boldsymbol{x},0)$ inside the integration over space defines an inner product in $L^2(\mathbb{R})$.

The terms in Eqs.(3.9) and (3.10) containing the medium matrices define the 'contrast-in-medium' contributions to the interaction of the two states. They vanish at those positions where $\delta^+_{QI}M^Y_{IP} - \delta^+_{PJ}M^Z_{JQ} = 0$, i.e., where $M^Y_{QP} = M^Z_{PQ}$. The terms containing the volume or surface source densities yield the contribution from the volume or surface sources, respectively, to the interaction of the two states. They vanish at source-free positions.

In a number of applications Eq.(3.10) will be applied to the entire $\mathbb{R}^3$. Then, outside some sphere $\mathcal{S}(\mathcal{O}, \Delta_0)$ with radius Δ_0 and center at the origin $\mathcal{O}$ of the chosen reference frame, the media in the two states will be assumed to be the same and homogenous as well as isotropic. For such a medium, the tensor Green's function is known analytically and in particular the causal and anti-causal source-type integral representations are known analytically. For the application of Eq.(3.10) to the entire $\mathbb{R}^3$, the theorem will be first applied to a sphere $\mathcal{S}(\mathcal{O}, \Delta)$ of radius Δ and center at the origin $\mathcal{O}$ of the chosen reference frame and the limit $\Delta \to \infty$ will be taken. If, now, in State Y the wavefield is causally related to the action of its volume or surface source distributions and in State Z the wavefield is anti-causally related to the action of its volume or surface source distributions (the volume or surface source distributions being assumed to have bounded supports), the integral over $\mathcal{S}(\mathcal{O}, \Delta)$ vanishes as $\Delta \to \infty$. However, if both states are causally related to the action of their volume sources, the integral over $\mathcal{S}(\mathcal{O}, \Delta)$ does not vanish as $\Delta \to \infty$, but has a constant value for sufficiently large values of Δ.

For State $Y =$ State Z and zero correlation time shift (i.e., $t = 0$), Eq.(3.9) reduces to the local energy balance for the wavefield and Eq.(3.10) to the global energy balance for the domain D provided that $M_{QP} = M_{PQ}$. This implies that for the energy considerations pertaining to a physical wavefield to hold, the medium matrix must be symmetric. In that case, also the quantity $\frac{1}{2}M_{PQ}F_PF_Q$ (whose time derivative occurs in Eqs.(3.9) and (3.10)) should represent the volume density of stored energy. For the latter, the symmetric medium matrix should, in addition, on physical grounds be positive definite.

Reciprocity property of the causal Green's tensor

Equation (3.8) leads to a reciprocity property of the Green's tensor which will be used in the wavefield extrapolation formulas to be discussed in Section 4. Let

$$F_P^Y = F_P^{Y;G}(\boldsymbol{x}, \boldsymbol{x}', t) \quad \text{with} \quad Q_I^Y = a_I^Y\, \delta_{\boldsymbol{x}'}(\boldsymbol{x})\delta_0(t)$$

be a causal wavefield in Medium Y. Then (cf. Eq.(2.28)), $F_P^{Y;G} = G_{PI}^Y(\boldsymbol{x}, \boldsymbol{x}', t)\, a_I^Y$. Let, similarly,

$$F_Q^Z = F_Q^{Z;G}(\boldsymbol{x}, \boldsymbol{x}'', t) \quad \text{with} \quad Q_J^Z = a_J^Z\, \delta_{\boldsymbol{x}''}(\boldsymbol{x})\delta_0(t)$$

be a causal wavefield in Medium Z. Then, $F_Q^{Z;G} = G_{QJ}^Z(\boldsymbol{x}, \boldsymbol{x}'', t)\, a_J^Z$. Take the media in the two states each other's adjoints, i.e., $M_{QP}^Y = M_{PQ}^Z$, and apply Eq.(3.8) to the entire $\mathbb{R}^3$. In this application, the contrast-in-media term and the contribution from the 'sphere at infinity' vanish. The result is

$$\delta_{\bar{Q}I}G_{QJ}^Z(\boldsymbol{x}', \boldsymbol{x}'', t)\, a_J^Z a_I^Y = \delta_{\bar{P}J}G_{PI}^Y(\boldsymbol{x}'', \boldsymbol{x}', t)\, a_I^Y a_J^Z \quad \text{for} \quad \boldsymbol{x}' \neq \boldsymbol{x}'' \,. \tag{3.11}$$

Since Eq.(3.11) has to hold for arbitrary values of a_I^Y and a_J^Z, we end up with

$$\delta_{\bar{Q}I}G_{QJ}^Z(\boldsymbol{x}', \boldsymbol{x}'', t) = \delta_{\bar{P}J}G_{PI}^Y(\boldsymbol{x}'', \boldsymbol{x}', t) \quad \text{for} \quad \boldsymbol{x}' \neq \boldsymbol{x}'' \,, \tag{3.12}$$

which entails

$$G_{IJ}^Z(\boldsymbol{x}', \boldsymbol{x}'', t) = G_{JI}^Y(\boldsymbol{x}'', \boldsymbol{x}', t) \quad \text{for} \quad \boldsymbol{x}' \neq \boldsymbol{x}'' \,. \tag{3.13}$$

Equation (3.13) is the reciprocity relation for the causal Green's tensor.

In the geometrical ray approximation, in accordance with Eq.(2.17) and $f_0(t) = \partial_t\delta(t)$, the causal Green's tensor can be written in the form

$$G_{PI'}(\boldsymbol{x}', \boldsymbol{x}'', t) \simeq \sum_{\text{modes,paths}} \Xi_P(\boldsymbol{x}')\Xi_{I'}(\boldsymbol{x}'')\, A^{(0)}(\boldsymbol{x}', \boldsymbol{x}'')\, \partial_t\delta(t - \tau(\boldsymbol{x}', \boldsymbol{x}'')) \,. \tag{3.14}$$

Since the reciprocity of the Green's tensor holds for any time, one expects reciprocity to hold at the arrival times of the direct waves, in particular. Indeed, it can be shown that the scalar amplitude must be symmetric in its arguments.

4 Wavefield extrapolation

In the reconstruction process, certain wavefield extrapolation formulas are needed to relate the different wavefields at arbitrary positions and times of observation to their respective volume or (actual or equivalent) surface source distributions. In these wavefield extrapolation formulas the causal Green's tensor for an unbounded domain occurs. The wavefield extrapolation formulas follow from the global reciprocity theorems given in Eqs.(3.8) and (3.10) by identifying State Y with the actual wavefield state (for which no superscript will be used) and State Z with a Green's state (to be denoted by the superscript G) that corresponds to a point-source excited causal wavefield defined in the entire $\mathbb{R}^3$. Then,

$$Q_J^G = a_J \, \delta_{\boldsymbol{x}'}(\boldsymbol{x})\delta_0(t) \ , \tag{4.1}$$

and

$$F_Q^G = G_{QJ}(\boldsymbol{x}, \boldsymbol{x}', t) \, a_J \ . \tag{4.2}$$

Further, we take

$$M_{PQ}^G = M_{QP} \quad \text{for} \quad \boldsymbol{x} \in D \ . \tag{4.3}$$

Application of Eq.(3.8) to the domain D then yields

$$\begin{aligned} &a_J \int_D \delta_{QI}^- \, \mathsf{C}_t(Q_I(\boldsymbol{x},.), G_{QJ}(\boldsymbol{x},\boldsymbol{x}',.)) \, \mathrm{d}V(\boldsymbol{x}) - a_J \, \delta_{PJ}^- \, F_P(\boldsymbol{x}',t) \, \chi_D(\boldsymbol{x}') \\ &= a_J \int_{\partial D} \delta_{QI}^- \mathcal{N}_{IP'} \, \mathsf{C}_t(F_{P'}(\boldsymbol{x},.), G_{QJ}(\boldsymbol{x},\boldsymbol{x}',.)) \, \mathrm{d}A(\boldsymbol{x}) \\ &+ a_J \int_S \delta_{QI}^- \, \mathsf{C}_t(q_I(\boldsymbol{x},.), G_{QJ}(\boldsymbol{x},\boldsymbol{x}',.)) \, \mathrm{d}A(\boldsymbol{x}) \quad \text{for} \quad \boldsymbol{x}' \in \mathbb{R}^3 \ . \end{aligned} \tag{4.4}$$

Using the reciprocity relation (3.12) for the causal Green's tensor, employing the fact that Eq.(4.4) has to hold for arbitrary values of a_J and observing that the factor δ_{PJ}^- is common to all terms, the result is

$$\begin{aligned} \chi_D(\boldsymbol{x}') \, F_P(\boldsymbol{x}',t) &= \int_D \mathsf{C}_t(Q_I(\boldsymbol{x},.), G_{PI}(\boldsymbol{x}',\boldsymbol{x},.)) \, \mathrm{d}V(\boldsymbol{x}) \\ &- \int_{\partial D} \mathcal{N}_{IP'} \, \mathsf{C}_t(F_{P'}(\boldsymbol{x},.), G_{PI}(\boldsymbol{x}',\boldsymbol{x},.)) \, \mathrm{d}A(\boldsymbol{x}) \\ &- \int_S \mathsf{C}_t(q_I(\boldsymbol{x},.), G_{PI}(\boldsymbol{x}',\boldsymbol{x},.)) \, \mathrm{d}A(\boldsymbol{x}) \quad \text{for} \quad \boldsymbol{x}' \in \mathbb{R}^3 \ . \end{aligned} \tag{4.5}$$

Equation (4.5) is the wavefield extrapolation formula of the time-convolution type. For $\boldsymbol{x}' \in D$ it is an integral representation for the wavefield that finds its application in the domain integral equation method for solving the scattering by penetrable objects. For $\boldsymbol{x}' \in \partial D$ it is used in the boundary integral equation method for solving scattering problems associated with the scattering by objects composed of constituting parts whose Green's tensor is explicitly known. For $\boldsymbol{x}' \in D'$ it represents Oseen's 'extinction theorem' and forms the basis for the 'null-field method' in solving scattering problems.

In an exactly similar manner, application of Eq.(3.10) to the domain D leads, upon taking again $M^G_{PQ} = M_{QP}$ for $\boldsymbol{x} \in D$, to

$$\chi_D(\boldsymbol{x}')\, F_P(\boldsymbol{x}',t) = \int_D \mathsf{R}_t(Q_I(\boldsymbol{x},.), G_{PI}(\boldsymbol{x}',\boldsymbol{x},.))\,\mathrm{d}V(\boldsymbol{x})$$

$$- \int_{\partial D} \mathcal{N}_{IP'}\, \mathsf{R}_t(F_{P'}(\boldsymbol{x},.), G_{PI}(\boldsymbol{x}',\boldsymbol{x},.))\,\mathrm{d}A(\boldsymbol{x})$$

$$- \int_S \mathsf{R}_t(q_I(\boldsymbol{x},.), G_{PI}(\boldsymbol{x}',\boldsymbol{x},.))\,\mathrm{d}A(\boldsymbol{x}) \quad \text{for} \quad \boldsymbol{x}' \in \mathbb{R}^3 \,. \tag{4.6}$$

Equation (4.6) is the wavefield extrapolation formula of the time-correlation type. Equations (4.5) and (4.6) find their use in seismic migration, viz., in the causal and anti-causal 'downward' continuation of the wavefield away from its physical or computational volume or surface sources.

5 Embedding procedure, contrast-source formulation and Born approximation

The first step towards constructing a solution of the inverse problem consists of applying an embedding procedure and introducing a contrast-source formulation. Subsequently, a linearization procedure, known as the Born approximation, is applied to the corresponding volume density of contrast source.

Embedding and contrast-source procedures

In the actual configuration, to be 'sensed', a target region of bounded support D_{con} is present where the constitutive properties of the medium differ by a certain amount from the ones of a known *embedding*. The embedding occupies the entire $\mathbb{R}^3$ and its Green's tensor is assumed to be known. Known volume or surface sources of fixed strengths and fixed bounded supports D_{src}

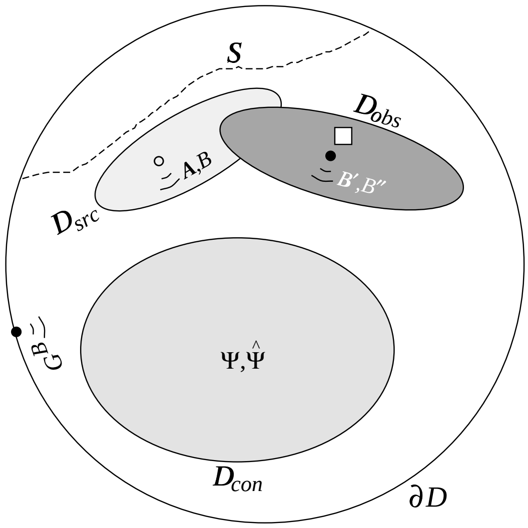

Figure 5.1: The configuration. (The blank box refers to Figure 6.1.)

and S_{src}, respectively, generate an interrogating wavefield. The difference between the values of this wavefield when generated in the embedding and the ones in the actual configuration are indicative for the presence of D_{con}. This difference wavefield is observed in a domain D_{obs} or on a surface S_{obs}, both of bounded support. The observations last a finite time interval $[0,T]$. In general, the intersection of D_{con} and either D_{obs} or S_{obs} is empty; this neednot be the case for either D_{src} and S_{src} on one hand and D_{obs} and S_{obs} on the other hand (see Figure 5.1). The wavefield state in the actual configuration will be denoted by the superscript A, the wavefield state in the embedding by the superscript B. Since $Q_I^A = Q_I^B$ and $q_I^A = q_I^B$, the difference $F_P^A - F_P^B$ between the wavefields F_P^A and F_P^B then satisfies the system of equations

$$(\mathcal{D}_{IP} + M_{IP}^B \partial_t)\,(F_P^A - F_P^B) = Q_I^{A,B}\ , \tag{5.1}$$

where

$$Q_I^{A,B} = -\mathcal{C}_{IP}^{A,B}\,\partial_t F_P^A\ , \tag{5.2}$$

with the contrast-in-medium properties

$$\mathcal{C}_{IP}^{A,B} = (M_{IP}^A - M_{IP}^B)\,\chi_{D_{con}}\ , \tag{5.3}$$

is the volume source density of the *contrast sources*, whose support is D_{con}, while $F_P^A - F_P^B$ is continuous across S_{src}.

In the actual computations, the quantities to be evaluated should be non-dimensionalized. For the principal quantity in our case, viz., the contrast-in-medium matrix, we write to this end

$$\mathcal{C}_{IP}^{A,B} = M_{IP'}^{B}\, \mathcal{X}_{P'P}^{A,B} \ ,$$

in which $\mathcal{X}_{P'P}^{A,B}$ is the 'relative' contrast-in-medium matrix.

The (first-order) Born approximation

For a given medium contrast $\mathcal{C}_{IP}^{A,B}$, the volume source density of contrast source would be known if the wavefield F_P^A occurring in the right-hand side of Eq.(5.2) were known. Evaluation of the latter would, however, require the solution of the relevant wavefield differential equation or its equivalent integral equation, which, in practice, goes at the cost of considerable numerical effort. For migration/inversion in remote sensing it is therefore common practice to apply an approximation procedure to the integral equation involved that consists of producing its Neumann expansion, of which expansion usually only the first-order term is retained. (For a time-domain convergence criterion of the relevant Neumann expansion, see De Hoop [27].) This first-order approximation is customarily referred to as the *(first-order) Born approximation*.

In the Born approximation, the unknown wavefield F_P^A in the right-hand side of Eq.(5.2) is replaced by its known zero-order approximation F_P^B for $\boldsymbol{x} \in D_{con}$. With this, we have

$$-Q_I^{A,B} = \mathcal{C}_{IP}^{A,B}\, \partial_t F_P^A \simeq \mathcal{C}_{IP}^{A,B}\, \partial_t F_P^B \ . \tag{5.4}$$

Note that the Born approximation only involves the volume contrast source density and that it does not affect the field in the chosen background medium of State B (that may be arbitrarily inhomogeneous and anisotropic). The use of the Born-approximated volume density of contrast source is also known as the *linearization procedure* in remote sensing. Considering the linearization as a variation, we may parametrize the medium M with parameters $\boldsymbol{\delta}$ that relate to particular properties of the medium. Then,

$$\mathcal{C}_{IP}^{A,B} \simeq \left.\frac{\partial(M_{IP})}{\partial(\boldsymbol{\delta}_\mu)}\right|_{\boldsymbol{\delta}^B} [\boldsymbol{\delta}^A - \boldsymbol{\delta}^B]_\mu \tag{5.5}$$

(summation over repeated Greek subscripts). For example, the parameters could be the medium's phase velocities which choice could improve the accuracy of the linearization for the transmitted scattered field.

Linearized reciprocity theorems of the time-convolution and the time-correlation types

In the construction of certain operators in the optimization approach for inversion, the linearized versions of the reciprocity theorems of the time-convolution and the time-correlation types are needed. In them one of the two states (State Y) is taken to be the state associated with the difference wavefield; the other is a computational state to be denoted by the superscript Z and which is to be determined later on. The theorems are applied to a domain D that contains D_{con} and D_{obs} in its interior.

With the use of Eq.(3.8), assuming that $M^Z_{PQ} = M^B_{QP}$, the linearized reciprocity theorem of the time-convolution type becomes (cf. Eq.(5.1))

$$\int_{D_{obs}} \delta^-_{PJ}\, \mathsf{C}_t(F^A_P - F^B_P, Q^Z_J)\,\mathrm{d}V(\boldsymbol{x}) \simeq$$

$$- \int_{D_{con}} \delta^-_{QI}\, \mathsf{C}_t(\mathcal{C}^{A,B}_{IP}\,\partial_t F^B_P, F^Z_Q)\,\mathrm{d}V(\boldsymbol{x})$$

$$- \int_{\partial D} \delta^-_{QI} \mathcal{N}_{IP}\, \mathsf{C}_t(F^A_P - F^B_P, F^Z_Q)\,\mathrm{d}A(\boldsymbol{x}) \; , \tag{5.6}$$

where we have confined the support of Q^Z_J to D_{obs} and have taken into account that the support of $Q^{A,B}_I$ is D_{con}.

In the case of *time-lapse* experiments, in this equation we would assign State A to the configuration at the 'new' time, State B to the initial or 'previous' configuration, and State Z to State B. An approach of this type ignores any non-linear changes in the constitutive parameters, in the elastic case for example due to changes in the stress field.

Similarly, with Eq.(3.10) and assuming that $M^Z_{PQ} = M^B_{QP}$, the linearized reciprocity theorem of the time-correlation type is obtained as

$$\int_{D_{obs}} \delta^+_{PJ}\, \mathsf{R}_t(F^A_P - F^B_P, Q^Z_J)\,\mathrm{d}V(\boldsymbol{x}) \simeq$$

$$\int_{D_{con}} \delta^+_{QI}\, \mathsf{R}_t(\mathcal{C}^{A,B}_{IP}\,\partial_t F^B_P, F^Z_Q)\,\mathrm{d}V(\boldsymbol{x})$$

$$+ \int_{\partial D} \delta^+_{QI} \mathcal{N}_{IP}\, \mathsf{R}_t(F^A_P - F^B_P, F^Z_Q)\,\mathrm{d}A(\boldsymbol{x}) \; . \tag{5.7}$$

In both theorems, the field $F^A_P - F^B_P$ is always causally related to its sources, because of the fact that F^A_P is associated with measured values of an actual, physical, wavefield. In the computational State Z, one is free to choose between a wavefield that is causally related to its sources and a wavefield that is anti-causally related to them. In the former case, the surface integral in Eq.(5.6) yields a vanishing contribution; in the latter case, the surface integral in Eq.(5.7) is zero.

Linearized wavefield extrapolation

The difference wavefield $F^A_P - F^B_P$ is expressible in terms of the volume density of its contrast source through the wavefield extrapolation formula of the time convolution type. Application of Eq.(4.5), with $\boldsymbol{x}$ and $\boldsymbol{x}'$ interchanged, to Eq.(5.1) and the entire $\mathbb{R}^3$ yields

$$(F^A_P - F^B_P)(\boldsymbol{x}, .) = \int_{D_{con}} \mathsf{C}_t(Q^{A,B}_I(\boldsymbol{x}', .), G^B_{PI}(\boldsymbol{x}, \boldsymbol{x}', .))\, \mathrm{d}V(\boldsymbol{x}')\ . \tag{5.8}$$

For the later use of this representation in its linearized version, it is notationally advantageous to consider it as an operator acting on the medium contrast. Correspondingly, we introduce the local kernel $\mathcal{E}^B_{PIP'}$ as

$$\mathcal{E}^B_{PIP'}(\boldsymbol{x}, \boldsymbol{x}', .) = \partial_t \mathsf{C}_t(F^B_{P'}(\boldsymbol{x}', .), G^B_{PI}(\boldsymbol{x}, \boldsymbol{x}', .)) \quad \text{on} \quad \mathbb{R}^3 \times D_{con}\ , \tag{5.9}$$

which entails

$$(F^A_P - F^B_P)(\boldsymbol{x}, .) \simeq -\int_{D_{con}} \mathcal{E}^B_{PIP'}(\boldsymbol{x}, \boldsymbol{x}', .)\, \mathcal{C}^{A,B}_{IP'}(\boldsymbol{x}')\, \mathrm{d}V(\boldsymbol{x}')\ , \tag{5.10}$$

and the global operator

$$\boldsymbol{L}^B_{PIP'}\ :\ \mathcal{C}^{A,B}_{IP'} \to (F^A_P - F^B_P)$$

(note the sign) through $F^A_P - F^B_P = \boldsymbol{L}^B_{PIP'}(\mathcal{C}^{A,B}_{IP'})$. The operator $\boldsymbol{L}^B$ can be interpreted as the Fréchet derivative of F^A with respect to M^A at M^B.

The *adjoint* of $\boldsymbol{L}^B$ with respect to the standard inner product in $L^2(\mathbb{R}^4)$ (over $D_{obs} \times \mathbb{R}$) follows from the linearized reciprocity theorem Eq.(5.7) with anti-causal State Z, taken at zero correlation time shift,

$$\int_{D_{obs}} \delta^+_{PJ}\, \mathsf{R}_t(\boldsymbol{L}^B_{PIP'}[\mathcal{C}^{A,B}_{IP'}], Q^Z_J)(\boldsymbol{x}, 0)\, \mathrm{d}V(\boldsymbol{x}) \simeq$$

$$\int_{D_{con}} \delta^+_{QI}\, \mathcal{C}^{A,B}_{IP}\ (\boldsymbol{L}^B)^T_{PQJ}[Q^Z_J]\, \mathrm{d}V(\boldsymbol{x}) \tag{5.11}$$

with

$$(\boldsymbol{L}^B)^T_{PQJ}[Q^Z_J](\boldsymbol{x}) = \partial_t \mathsf{R}_t(F^B_P, F^Z_Q)(\boldsymbol{x}, 0)\ . \tag{5.12}$$

In case the State Z is chosen to be causal, a boundary contribution to the adjoint occurs. Consider the boundary integral of Eq.(5.7), substitute Eq.(5.10) and interchange the order of integration,

$$\int_{\partial D} \delta^+_{QJ} \mathcal{N}_{JP}(\boldsymbol{x})\, \mathsf{R}_t(\boldsymbol{L}^B_{PIP'}[\mathcal{C}^{A,B}_{IP'}], F^Z_Q)(\boldsymbol{x}, 0)\, \mathrm{d}A(\boldsymbol{x}) = \tag{5.13}$$

$$- \int_{D_{con}} \int_{\partial D} \delta^+_{QJ} \mathcal{N}_{JP}(\boldsymbol{x})\, \mathsf{R}_t(\mathcal{E}^B_{PIP'}(\boldsymbol{x}, \boldsymbol{x}', .), F^Z_Q(\boldsymbol{x}, .))(0)\, \mathrm{d}A(\boldsymbol{x})\, \mathcal{C}^{A,B}_{IP'}(\boldsymbol{x}')\, \mathrm{d}V(\boldsymbol{x}') \ .$$

Thus, the *complete* adjoint of $\boldsymbol{L}^B$ can be written as

$$(\boldsymbol{L}^B)^T_{PQJ}[Q^Z_J](\boldsymbol{x}) = \partial_t \mathsf{R}_t(F^B_P, F^Z_Q)(\boldsymbol{x}, 0) \tag{5.14}$$

$$- \int_{\partial D} \mathcal{N}_{IP'}(\boldsymbol{x}')\, \mathsf{R}_t(\mathcal{E}^B_{P'QP}(\boldsymbol{x}', \boldsymbol{x}, .), F^Z_I(\boldsymbol{x}', .))(0)\, \mathrm{d}A(\boldsymbol{x}') \ .$$

Linearized wavefield extrapolation in the geometrical ray approximation

In this subsection, we complement the previous analysis with an analysis of the singular behavior of the linearized wavefield extrapolator and its adjoint. To this end, we will constrain ourselves to a point source of fixed type I', i.e.,

$$Q^B_J = \delta^+_{JI'} \delta_{\boldsymbol{s}}(\boldsymbol{x}) \delta_0(t) \ ,$$

indicating the source location by $\boldsymbol{s}$. Let the point of observation be $\boldsymbol{x} = \boldsymbol{r}$. Then, in $\mathcal{E}^B_{PIP'}$, using reciprocity we have

$$F^B_{P'}(\boldsymbol{x}', .) = G^B_{P'I'}(\boldsymbol{x}', \boldsymbol{s}, .) \ , \quad G^B_{PI}(\boldsymbol{x}, \boldsymbol{x}', .) = G^B_{IP}(\boldsymbol{x}', \boldsymbol{r}, .) \ . \tag{5.15}$$

To show whether certain quantities relate to the first ($\boldsymbol{s}$) or the second ($\boldsymbol{r}$) field in the equation above, we employ $\tilde{.}$'s for the first and $\hat{.}$'s for the second. We will omit the superscript B for both fields. While applying the geometrical ray approximation, we will separate the contributions from the different modes and paths, and consider these independently. Thus, in the geometrical ray approximation Eq.(5.9) becomes

$$\mathcal{E}^B_{PIP'}(\boldsymbol{r}, \boldsymbol{x}', .) \simeq \tilde{\Xi}_{I'}(\boldsymbol{s}) \hat{\Xi}_P(\boldsymbol{r}) \tilde{\Xi}_{P'}(\boldsymbol{x}') \hat{\Xi}_I(\boldsymbol{x}')\, \tilde{A}^{(0)}(\boldsymbol{x}') \hat{A}^{(0)}(\boldsymbol{x}')\, \partial^3_t \delta(t - \tilde{\tau}(\boldsymbol{x}') - \hat{\tau}(\boldsymbol{x}')) \ . \tag{5.16}$$

The integral over D_{con} in Eq.(5.10) reduces to an integral over the *isochron*

$$\{\boldsymbol{x}' \in D_{con} \mid \tilde{\tau}(\boldsymbol{x}') + \hat{\tau}(\boldsymbol{x}') = t\} \ , \tag{5.17}$$

and thus obtains the structure of a Generalized Radon Transform. The properties of this particular transform have been discussed in detail by Rakesh [28].

To analyze the adjoint extrapolation operator, let Q^Z_J be represented by

$$Q^Z_J(\boldsymbol{x},t) = \int \delta^+_{JP'} \delta_{\boldsymbol{r}}(\boldsymbol{x})\, Q^Z_{P'}(\boldsymbol{r},t)\,\mathrm{d}\boldsymbol{r} \ .$$

Then Eq.(5.12) with the aid of eq.(3.13) leads to

$$(\boldsymbol{L}^B)^T_{PQJ}[Q^Z_J](\boldsymbol{x}) = -\int \mathsf{R}_t(G^B_{I'P}(\boldsymbol{s},\boldsymbol{x},.), \mathsf{C}_t(G^B_{JQ}(\boldsymbol{r},\boldsymbol{x},.), \partial_t Q^Z_J(\boldsymbol{r},.)))(0)\,\mathrm{d}\boldsymbol{r} \ ; \tag{5.18}$$

in the geometrical ray approximation the right-hand side becomes

$$(\boldsymbol{L}^B)^T_{PQJ}[Q^Z_J](\boldsymbol{x}) \simeq \tag{5.19}$$

$$\int \tilde{\Xi}_{I'}(\boldsymbol{s})\hat{\Xi}_J(\boldsymbol{r})\tilde{\Xi}_P(\boldsymbol{x})\hat{\Xi}_Q(\boldsymbol{x})\,\tilde{A}^{(0)}(\boldsymbol{x})\hat{A}^{(0)}(\boldsymbol{x})\,\partial^3_t Q^Z_J(\boldsymbol{r},\tilde{\tau}(\boldsymbol{x})+\hat{\tau}(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{r} \ ,$$

which is an integration over the *diffraction set* (Hagedoorn [29])

$$\{(\boldsymbol{r},t) \in D_{obs} \times [0,T] \mid t = \tilde{\tau}(\boldsymbol{x}) + \hat{\tau}(\boldsymbol{x})\} \ , \tag{5.20}$$

and has in fact the structure of the dual Generalized Radon Transform. Beylkin [30] introduced this dual to derive the inverse Generalized Radon Transform. For the geometrical ray approximations of the wavefield extrapolator and its adjoint in State B, we will employ the shorthand notation $\overset{\infty}{\boldsymbol{L}}_{PIP'}$ and $(\overset{\infty}{\boldsymbol{L}})^T_{PQJ}$, respectively.

Equation (5.18) can be given a processing-sequential interpretation. Let the source Q^Z_J depend on $\boldsymbol{s}$ also, such that we can substitute here measurements in the later analysis. Then, the time-reversed convolution of $G^B_{JQ}(\boldsymbol{r},\boldsymbol{x},.)$ with $Q^Z_J(\boldsymbol{r},\boldsymbol{s},.)$ represents a *depropagation* procedure. Carrying out the integration over $\boldsymbol{r}$, keeping $\boldsymbol{x}$ fixed, results in a so-called Common Focal Point (CFP) gather, representing the outcome of the integration as function of $\boldsymbol{s}$ and t. The second *focussing* step yields the time-convolution of $G^B_{I'P}(\boldsymbol{s},\boldsymbol{x},.)$ with the CFP gather followed by an integration over $\boldsymbol{s}$, see Berkhout and Rietveld [31] and Thorbecke [32].

6 The local optimization approach to inversion

In the inverse scattering problem we will distinguish two aspects. One aspect, discussed in this section, concerns the reconstruction or ‘imaging’ of a small perturbation superimposed on a given background medium. The second aspect, addressed in the next section, concerns improving the background with the outcome of the local optimization procedure.

The optimization approach to inversion starts with selecting a suitable error criterion that quantifies the 'misfit' $\epsilon^{A,B}$ between the wavefields in the States A and B as far as they are observed in the observational domain D_{obs}. Let the measuring process be described by the linear operation $W_{MP}(\sigma)$, defined on D_{obs} for all t, and parametrized by σ. A choice of misfit that suits our purpose is provided by

$$\epsilon^{A,B}(\sigma) = \int_{D_{obs}} \int_{t\in\mathbb{R}} [W_{MP}(\sigma)(F_P^A - F_P^B)]\, \delta^+_{MN}\, [W_{NQ}(\sigma)(F_Q^A - F_Q^B)]\, \mathrm{d}V(\boldsymbol{x})\, \mathrm{d}t \ . \qquad (6.1)$$

The operator $W_{MP}(\sigma)$ may reveal the *statistical* variations in the measurements. Obviously, $\epsilon^{A,B} > 0$ when $F_P^B \not\equiv F_P^A$ for $\boldsymbol{x} \in D_{obs}$. The expression for $\epsilon^{A,B}$ can be regarded as a time correlation at zero time shift, viz.

$$\epsilon^{A,B}(\sigma) = \int_{D_{obs}} \delta^+_{MN}\, \mathsf{R}_t(W_{MP}(\sigma)(F_P^A - F_P^B), W_{NQ}(\sigma)(F_Q^A - F_Q^B))(\boldsymbol{x},0)\, \mathrm{d}V(\boldsymbol{x}) \ . \qquad (6.2)$$

Introducing, under appropriate conditions, the adjoint W^T of the operator W with respect to the real inner product on $L^2(D_{obs} \times \mathbb{R})$, the latter equation can be written in the form

$$\epsilon^{A,B}(\sigma) = \int_{D_{obs}} \delta^+_{PJ}\, \mathsf{R}_t((F_P^A - F_P^B), (W^TW)_{JQ}(\sigma)(F_Q^A - F_Q^B))(\boldsymbol{x},0)\, \mathrm{d}V(\boldsymbol{x}) \ . \qquad (6.3)$$

Note that W^TW is self-adjoint and positive, and can really be considered as a single operator or pre-processor. In view of the reciprocity relations to apply, it should satisfy the criterion that

$$(W^TW)_{JQ}(\sigma)(F_Q^A - F_Q^B)$$

defines an equivalent source distribution, in a configuration with the medium parameters of State B. A crude choice that satisfies this criterion is

$$(W^TW)_{JQ}(\sigma) = M^B_{JQ}/T \ ,$$

in which T is the measuring time interval as before. Probabilistic inversion approaches incorporating *a posteriori* statistics of the observations, typically replace M^B by a correlation matrix function with the assumption of Gaussian statistics.

Often, in practice, W^TW will be represented by a surface integral

$$(W^TW)_{JQ}(\sigma) = \chi_{[0,T]}(.) \int_{S_{obs}} \mathrm{d}\boldsymbol{r}\, S_{JQ}(\boldsymbol{r}, .; \sigma)\, \delta_{\boldsymbol{r}} \ , \quad S_{obs} \subset D_{obs} \ ,$$

or a sum of Dirac distributions,

$$(W^TW)_{JQ}(\sigma) = \chi_{[0,T]}(.)\, \frac{|S_{obs}|}{N_{obs}} \sum_{[n]=1}^{N_{obs}} S_{JQ}(\boldsymbol{r}[n], .; \sigma)\, \delta_{\boldsymbol{r}[n]} \ , \quad \boldsymbol{r}[n] \in D_{obs} \ ,$$

which turns the misfit into an integral or a sum over point-receiver locations.

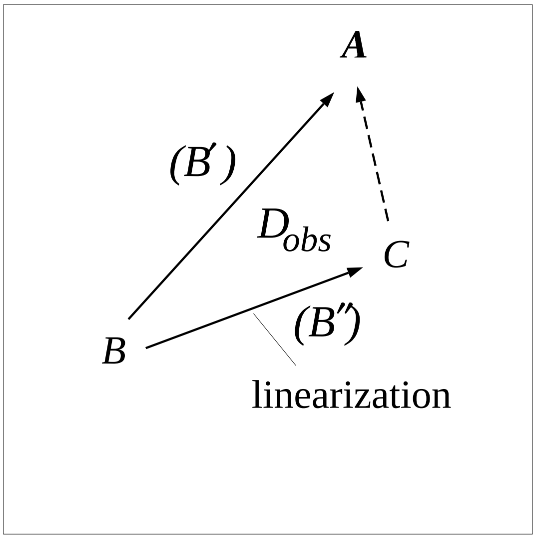

Figure 6.1: The optimization states.

The optimization condition

The aim of the optimization approach is to construct a State C with a medium matrix M^C_{IP} that deviates from M^B_{IP} only in the domain D_{con}, has an accompanying causal wavefield F^C_P that is generated by the same sources as in the States A and B, i.e., $Q^C_I = Q^B_I = Q^A_I$ and $q^C_I = q^B_I = q^A_I$, and that has the property that the mismatch between the States A and C is less than the mismatch between the States A and B. Now, with $F^A_P - F^B_P = (F^A_P - F^C_P) + (F^C_P - F^B_P)$, Eq.(6.2) can be rewritten as

$$\epsilon^{A,B} = \epsilon^{A,C} + 2\int_{D_{obs}}\int_{t\in\mathbb{R}} \delta^+_{PJ}\,(F^C_P - F^B_P)\,(W^TW)_{JQ}(F^A_Q - F^C_Q)\,\mathrm{d}V(\boldsymbol{x})\,\mathrm{d}t$$
$$+\,\epsilon^{B,C}\ . \tag{6.4}$$

We assume that $F^C_P \not\equiv F^B_P$ for $\boldsymbol{x} \in D_{obs}$ i.e. the contrast volume source density associated with the medium update $\mathcal{C}^{C,B}_{IP}$ in D_{con} is assumed not to lead to a vanishing difference field in D_{obs}. Then, $\epsilon^{B,C} > 0$. As a consequence of Eq.(6.4), the condition $\epsilon^{A,C} < \epsilon^{A,B}$ is met by requiring that

$$\int_{D_{obs}}\int_{t\in\mathbb{R}} \delta^+_{PJ}\,(F^C_P - F^B_P)\,(W^TW)_{JQ}(F^A_Q - F^C_Q)\,\mathrm{d}V(\boldsymbol{x})\,\mathrm{d}t = 0\ , \tag{6.5}$$

which is the *optimization condition*. To rewrite it in a form in which the difference between data and simulation $F^A_Q - F^B_Q$ occurs, we substitute

$$F^A_Q - F^C_Q = (F^A_Q - F^B_Q) - (F^C_Q - F^B_Q) \ ,$$

and obtain

$$\int_{D_{obs}} \int_{t\in\mathbb{R}} \delta^{+}_{PJ} \, (F^C_P - F^B_P) \, (W^T W)_{JQ} (F^A_Q - F^B_Q) \, \mathrm{d}V(\boldsymbol{x}) \, \mathrm{d}t$$
$$= \int_{D_{obs}} \int_{t\in\mathbb{R}} \delta^{+}_{PJ} \, (F^C_P - F^B_P) \, (W^T W)_{JQ} (F^C_Q - F^B_Q) \, \mathrm{d}V(\boldsymbol{x}) \, \mathrm{d}t \ . \tag{6.6}$$

Note that the right-hand side of this equation is positive, thus the left-hand side has to be positive as well. From it we have to determine $\mathcal{C}^{C,B}_{IP}$ with its support D_{con}. To relate the difference field in D_{obs} to its contrast sources in D_{con} yet to be determined, the reciprocity relation of the time correlation type is applied, first, to the left-hand side. To this end, we consider the quantity $F^B_P - F^C_P$ as a wavefield and the quantity

$$Q^{B'}_J = (W^T W)_{JQ} (F^A_Q - F^B_Q) \tag{6.7}$$

as the (known) volume source density with D_{obs} as its support, of a computational State B' with associated (causal or anti-causal) wavefield $F^{B'}_P$ in the medium with the known medium matrix

$$M^{B'}_{PQ} = M^B_{QP} \ .$$

Then, application of Eq.(3.10) to the domain interior to a closed surface ∂D that completely surrounds both D_{obs} and D_{con}, i.e., $D_{obs} \cup D_{con} \subset D$, yields

$$\int_{D_{obs}} \int_{t\in\mathbb{R}} \delta^{+}_{PJ} \, (F^C_P - F^B_P) \, (W^T W)_{JQ} (F^A_Q - F^B_Q) \, \mathrm{d}V(\boldsymbol{x}) \, \mathrm{d}t$$
$$= \int_{D_{obs}} \delta^{+}_{PJ} \, \mathsf{R}_t (F^C_P - F^B_P, (W^T W)_{JQ} (F^A_Q - F^B_Q))(\boldsymbol{x}, 0) \, \mathrm{d}V(\boldsymbol{x})$$
$$= \int_{D_{con}} \delta^{+}_{QI} \, \mathsf{R}_t (\partial_t F^C_P, F^{B'}_Q)(\boldsymbol{x}, 0) \, \mathcal{C}^{C,B}_{IP}(\boldsymbol{x}) \, \mathrm{d}V(\boldsymbol{x})$$
$$+ \int_{\partial D} \delta^{+}_{QI} \mathcal{N}_{IP} \, \mathsf{R}_t (F^C_P - F^B_P, F^{B'}_Q)(\boldsymbol{x}, 0) \, \mathrm{d}A(\boldsymbol{x}) \ . \tag{6.8}$$

Carrying out a similar procedure to the right-hand side of Eq.(6.6), involves the introduction of a computational State B'' with volume source density

$$Q^{B''}_J = (W^T W)_{JQ} (F^C_Q - F^B_Q) \tag{6.9}$$

and medium matrix

$$M^{B''}_{PQ} = M^{B'}_{PQ} \; .$$

In the latter derivation, replace $(F^A_Q - F^B_Q)$ by $(F^C_Q - F^B_Q)$ and $F^{B'}_Q$ by $F^{B''}_Q$ (see Figure 6.1 for an illustration of the different states). Substituting the result and Eq.(6.8) into Eq.(6.6), the optimization condition takes the form

$$\mathcal{P} = \mathcal{Q} \; , \tag{6.10}$$

in which

$$\begin{aligned}\mathcal{P} = & \int_{D_{con}} \delta^+_{QI} \, \mathsf{R}_t(\partial_t F^C_P, F^{B'}_Q)(\boldsymbol{x},0) \, \mathcal{C}^{C,B}_{IP}(\boldsymbol{x}) \, \mathrm{d}V(\boldsymbol{x}) \\ & + \int_{\partial D} \delta^+_{QI} \, \mathcal{N}_{IP} \, \mathsf{R}_t(F^C_P - F^B_P, F^{B'}_Q)(\boldsymbol{x},0) \, \mathrm{d}A(\boldsymbol{x}) \; , \\ \mathcal{Q} = & \int_{D_{con}} \delta^+_{QI} \, \mathsf{R}_t(\partial_t F^C_P, F^{B''}_Q)(\boldsymbol{x},0) \, \mathcal{C}^{C,B}_{IP}(\boldsymbol{x}) \, \mathrm{d}V(\boldsymbol{x}) \\ & + \int_{\partial D} \delta^+_{QI} \, \mathcal{N}_{IP} \, \mathsf{R}_t(F^C_P - F^B_P, F^{B''}_Q)(\boldsymbol{x},0) \, \mathrm{d}A(\boldsymbol{x}) = \epsilon^{C,B} \; . \end{aligned} \tag{6.11}$$

In these expressions, the wavefield $F^C_P - F^B_P$ satisfies the equation (cf. Eq.(5.1))

$$(\mathcal{D}_{IP} + M^B_{IP}\partial_t)\,(F^C_P - F^B_P) = Q^{C,B}_I \; ,$$

where (cf. Eq.(5.2))

$$Q^{C,B}_I = -\mathcal{C}^{C,B}_{IP} \, \partial_t F^C_P \; .$$

So far, no approximations have been made. However, since $\mathcal{C}^{C,B}_{IP}$ is yet to be determined, we are inclined to avoid the evaluation of F^C. To this end, we *locally linearize* the expressions above in the contrast sources. This means that

- in the volume integrals F^C_P is replaced by F^B_P (cf. Eq.(5.4)),
- in the surface integrals $F^C_P - F^B_P$ is replaced by $\boldsymbol{L}^B_{PIP'}[\mathcal{C}^{C,B}_{IP'}]$ (cf. Eq.(5.10)),
- the source distribution $Q^{B''}_J$ is replaced by $(W^T W)_{JQ} \boldsymbol{L}^B_{QIP'}[\mathcal{C}^{C,B}_{IP'}]$ (cf. Eq.(5.10)).

We denote the linearized versions of $\mathcal{P}$ and $\mathcal{Q}$ by $\mathcal{P}_{lin}$ and $\mathcal{Q}_{lin}$, respectively.

For the wavefield extrapolation in states B' and B'' we have the choice between a causal and an anti-causal one. For the anti-causal extrapolation the integral over ∂D in $\mathcal{P}$ and $\mathcal{Q}$ vanishes.

Imaging and the improvement condition

Into the boundary integrals occurring in the expressions for $\mathcal{P}_{lin}$ and $\mathcal{Q}_{lin}$ we will now substitute Eq.(5.10). As in the derivation of Eq.(5.14) we will interchange the integrations over ∂D and over D_{con}. Then

$$\mathcal{P}_{lin} = \int_{D_{con}} \delta^{+}_{QI}\, \Psi^{B'}_{PQ}(\boldsymbol{x})\, \mathcal{C}^{C,B}_{IP}(\boldsymbol{x})\, \mathrm{d}V(\boldsymbol{x}) \tag{6.12}$$

with

$$\Psi^{B'}_{PQ}(\boldsymbol{x}) = \partial_t \mathsf{R}_t(F^{B}_{P}, F^{B'}_{Q})(\boldsymbol{x}, 0) - \int_{\partial D} \mathcal{N}_{I'P'}(\boldsymbol{x}')\, \mathsf{R}_t(\mathcal{E}^{B}_{P'QP}(\boldsymbol{x}', \boldsymbol{x}, .), F^{B'}_{I'}(\boldsymbol{x}', .))(0)\, \mathrm{d}A(\boldsymbol{x}') \ . \tag{6.13}$$

In this expression we identify two kernels,

$$\mathcal{I}^{B,B'}_{PQ}(\boldsymbol{x}) = \partial_t \mathsf{R}_t(F^{B}_{P}, F^{B'}_{Q})(\boldsymbol{x}, 0) \quad \text{on} \quad D_{con} \tag{6.14}$$

and

$$\mathcal{B}^{B,B'}_{PQ}(\boldsymbol{x}, \boldsymbol{x}') = \mathcal{N}_{I'P'}(\boldsymbol{x})\, \mathsf{R}_t(\mathcal{E}^{B}_{P'QP}(\boldsymbol{x}', \boldsymbol{x}, .), F^{B'}_{I'}(\boldsymbol{x}, .))(0) \quad \text{on} \quad D_{con} \times \partial D \ . \tag{6.15}$$

A likewise procedure applied to $\mathcal{Q}_{lin}$ leads to the introduction of $\Psi^{B''}_{PQ}(\boldsymbol{x})$. Unlike $\Psi^{B'}_{PQ}(\boldsymbol{x})$, the matrix $\Psi^{B''}_{PQ}(\boldsymbol{x})$ depends on the medium update $\mathcal{C}^{C,B}_{IP}$ through the volume source density of State B''. The presence of concatenation matrix δ^{+}_{QI} in Eq.(6.12) amounts to taking the *trace* of the integrand. The kernel $\mathcal{I}^{B,B'}$ can be considered as a L^2-inner product with respect to time evaluated at each point in D_{con}. It is also noticed that this kernel is a function of the actual source and receiver locations associated with our single experiment.

To construct an approximate solution, the unknown matrix function $\mathcal{C}^{C,B}$ is now written as

$$\mathcal{C}^{C,B}_{IP}(\boldsymbol{x}) = \alpha\, \Phi_{IP}(\boldsymbol{x}) \ , \tag{6.16}$$

where $\alpha = \alpha(\sigma)$ is an expansion coefficient and Φ is an expansion function belonging to the same space as $\mathcal{C}^{C,B}$ (i.e., it is supposed to have the same structure and have its support in D_{con}). In applications, Φ is referred to as the *image* matrix. In accordance with the local linearization and Eq.(6.16), we replace State B'' by a State $\widehat{B}$ with

$$Q^{B''}_{J} = \alpha\, Q^{\widehat{B}}_{J} \ \text{ hence } \ F^{B''}_{Q} = \alpha\, F^{\widehat{B}}_{Q} \ , \tag{6.17}$$

where $F^{\widehat{B}}$ is the wavefield extrapolated away from $Q^{\widehat{B}}$. On the basis of expansion Eq.(6.16), we introduce $\widehat{\mathcal{P}}$ and $\widehat{\mathcal{Q}}$ as

$$\mathcal{P}_{lin} = \alpha\,\widehat{\mathcal{P}} \quad \text{and} \quad \mathcal{Q}_{lin} = \alpha^2\,\widehat{\mathcal{Q}}\ . \tag{6.18}$$

Then solving condition (6.10) amounts to

$$\alpha = \widehat{\mathcal{P}}/\widehat{\mathcal{Q}}\ . \tag{6.19}$$

With State $\widehat{B}$ is associated a matrix function $\Psi^{\widehat{B}}(\boldsymbol{x})$ defined through Eq.(6.13). Thus we find

$$\widehat{\mathcal{P}} = \int_{D_{con}} \delta^{+}_{QI}\,\Psi^{B'}_{PQ}(\boldsymbol{x})\Phi_{IP}(\boldsymbol{x})\,\mathrm{d}V(\boldsymbol{x})\ , \tag{6.20}$$

$$\widehat{\mathcal{Q}} = \int_{D_{con}} \delta^{+}_{QI}\,\Psi^{\widehat{B}}_{PQ}(\boldsymbol{x})\Phi_{IP}(\boldsymbol{x})\,\mathrm{d}V(\boldsymbol{x})\ . \tag{6.21}$$

To achieve anything at all, viz., $\alpha \neq 0$, the *improvement condition* $\widehat{\mathcal{P}} \neq 0$, must be satisfied. (Note, however, that this condition does not guarantee convergence of the procedure.) To achieve positive definiteness of the entire domain integral, as in the gradient method, we choose

$$\Phi_{IP}(\boldsymbol{x}) = \mathrm{Op}\left[\delta^{+}_{QI}\,\Psi^{B'}_{PQ}(\boldsymbol{x}) + \eta\,\phi_{IP}\right]\chi_{D_{con}}(\boldsymbol{x})\ ; \tag{6.22}$$

η is a scalar quantity, and the matrix ϕ_{IP} is constant and equal to

$$\phi_{IP} = \frac{1}{|D_{con}|}\int_{D_{con}} \delta^{+}_{QI}\,\Psi^{B'}_{PQ}(\boldsymbol{x})\,\mathrm{d}V(\boldsymbol{x})\ , \tag{6.23}$$

the volume average of $\Psi^{B'}_{PI}$. The operator Op can be employed to control the *scale* or regularity of medium variations; in the gradient method Op simply equals the identity. The improvement condition requires that

$$\eta \neq -\frac{\dfrac{1}{|D_{con}|}\displaystyle\int_{D_{con}} \left(\delta^{+}_{QI}\,\Psi^{B'}_{PQ}(\boldsymbol{x})\right)^2 \mathrm{d}V(\boldsymbol{x})}{\left(\dfrac{1}{|D_{con}|}\displaystyle\int_{D_{con}} \delta^{+}_{QI}\,\Psi^{B'}_{PQ}(\boldsymbol{x})\,\mathrm{d}V(\boldsymbol{x})\right)^2} \tag{6.24}$$

(summed over I's and P's). If either $\widehat{\mathcal{P}}$ or $\widehat{\mathcal{Q}}$ becomes zero, the scheme fails unless we are, for $\widehat{\mathcal{P}} = 0$, at the exact solution already. It is observed that

- $\mathcal{I}^{B,B'}$ corresponds with the image from a 'pre-stack seismic depth migration' [14, 33, 34] if F^B is neglected with respect to F^A in $Q^{B'}$ (this requires the medium of State B to be sufficiently smooth);

- the boundary integral of $\mathcal{B}^{B,B'}$ contains the computational states only, hence measurements on the boundary are not required;

- the contribution from the constant matrix ϕ plays an important rôle where headwave or critical angle phenomena are considered for inversion ($\eta = -1$);

- the scheme allows for 'multi-pathing' and caustics in the computational states;

- an iterative procedure of the steepest-descent type follows from the identification of State C with the $(m+1)$-st step and of State B with the m-th step; the procedure is terminated whenever the mismatch $\epsilon^{A,B}(\sigma)$ is below a prescribed threshold.

In the process of contructing the expansion function, we are not bound by the precise choice of Eq.(6.22). As long as the improvement condition is satisfied, we can adjust this particular choice. Numerical experiments have indicated that invoking a multi-resolution analysis for $\Psi^{B'}$, i.e., allowing finer scales gradually with the number of iterations, may stabilize the inversion [35]. Also, we are not bound by the L^2 norm; how to change the procedure to other norms is discussed in Appendix C.

To evaluate (propagate away from the source) F^B one can employ Eq.(4.5). Taking the field $F^{B'}$ to be the anti-causal one, which implies that the boundary integrals vanish, extrapolation (away from the observations) can be based on Eq.(4.6).

The 'missing' field

The field $F^{B'}$ can be interpreted as the 'missing' field in the configuration, generated by the observed difference field $F^A - F^B$, and extrapolated away ('backpropagated') from the observations in medium M^B in the absence of the contrast source distribution. Hence the fields $F^{B'}$ and $F^A - F^B$ will differ from one another, in particular due to the scattering from the domain D_{con}. The optimization approach, basically, tries to minimize the missing field.

We will discuss a way to rewrite $F^{B'}$ in the idealized case where point observations have been carried out on ∂D, i.e., all around the domain D_{con}. Thus

$$F^{B'}|_{\partial D} = (F^A - F^B)|_{\partial D} \ .$$

Then, in D_{con}, the missing field can be expressed in the actual field F^A. Using Eq.(4.6) (source-free) in $D \backslash D_{obs}$ we get

$$\chi_D(\boldsymbol{x})\, F_P^{B'}(\boldsymbol{x}, t) = \int_{\partial D} \mathcal{N}_{IP'}(\boldsymbol{x}')\, \mathsf{R}_t((F_{P'}^A - F_{P'}^B)(\boldsymbol{x}', .)\, , G_{PI}^B(\boldsymbol{x}', \boldsymbol{x}, .))\, \mathrm{d}A(\boldsymbol{x}')\ .$$

Applying the reciprocity relation (3.10) to rewrite the latter surface integral representation, amounts to

$$\chi_D(\boldsymbol{x})\, F_P^{B'}(\boldsymbol{x}, t) = (F_P^A - F_P^B)(\boldsymbol{x}, t) + \int_D \mathsf{R}_t(Q_I^{A,B}(\boldsymbol{x}', .)\, , G_{PI}^B(\boldsymbol{x}', \boldsymbol{x}, .))\, \mathrm{d}V(\boldsymbol{x}')$$

$$= \int_{D_{con}} \mathsf{C}_t(Q_I^{A,B}(\boldsymbol{x}', .)\, , G_{IP}^B(\boldsymbol{x}, \boldsymbol{x}', .) + \mathsf{T} G_{PI}^B(\boldsymbol{x}', \boldsymbol{x}, .))\, \mathrm{d}V(\boldsymbol{x}')\ . \quad (6.25)$$

Here, with Eq.(5.2),

$$Q_I^{A,B} = -\mathcal{C}_{IP'}^{A,B}\, \partial_t F_{P'}^A\ .$$

Operator formalism

We reconsider the expressions for $\Psi_{PQ}^{B'}$, $\Psi_{PQ}^{B''}$ and $\Psi_{PQ}^{\widehat{B}}$ and write them in terms of the operator $\boldsymbol{L}^B$ introduced in Section 5. Using Eq.(5.14), we find that

$$\Psi_{PQ}^{B'} = (\boldsymbol{L}^B)_{PQJ}^T[(W^T W)_{JQ'}(F_{Q'}^A - F_{Q'}^B)]\ , \quad (6.26)$$

while, in conjunction with Eq.(5.10), we have

$$\Psi_{PQ}^{B''} = (\boldsymbol{L}^B)_{PQJ}^T[(W^T W)_{JQ'}\boldsymbol{L}_{Q'IP'}^B[\mathcal{C}_{IP'}^{C,B}]]\ , \quad (6.27)$$

and

$$\Psi_{PQ}^{\widehat{B}} = (\boldsymbol{L}^B)_{PQJ}^T[(W^T W)_{JQ'}\boldsymbol{L}_{Q'IP'}^B[\Phi_{IP'}]]\ . \quad (6.28)$$

The operator formalism is convenient to expose the connection of the proposed iterative inversion procedure with imaging and direct, weak, linearized inversion procedures. The main thing to note is that according to Eq.(5.19) in the geometrical ray approximation $\Psi_{PQ}^{B'}$ is the outcome of a ‘weighted diffraction stack’, which is represented in wave-theoretical form by Eq.(5.18). In this context $\boldsymbol{x} \in D_{con}$ is referred to as the (geometrical) image or the (wave) focal point, respectively.

7 The method of preconditioning

Weak, linearized inversion

In the steepest-descent-type iterative inversion scheme of Section 6, where $-\Phi_{IP'}$ is the 'gradient', the improvement condition led to medium update,

$$\Phi_{IP'} = \Psi^{B'}_{P'I} \; , \quad \mathcal{C}^{C,B}_{IP'} = \alpha \, \Phi_{IP'} = \alpha \, \Psi^{B'}_{P'I} \quad \text{with} \quad \alpha = \widehat{\mathcal{P}} / \widehat{\mathcal{Q}} \; .$$

Here, α is just a multiplicative factor, which through σ may vary with the image point. In the direct weak, linearized inversion scheme on the other hand, we would have to solve the matrix integral equation

$$\Psi^{B''}_{PQ} = \Psi^{B'}_{PQ} \quad \text{with} \quad \Psi^{B''}_{PQ} = (\boldsymbol{L}^B)^T_{PQJ} [(W^T W)_{JQ'} \boldsymbol{L}^B_{Q'IP'} [\mathcal{C}^{A,B}_{IP'}]] \; , \tag{7.1}$$

in accordance with Eq.(6.27). Under certain constraints, $(\boldsymbol{L}^B)^T_{PQJ} \; (W^T W)_{JQ'} \boldsymbol{L}^B_{Q'IP'}$ is a pseudo-differential operator and its *parametrix* exists. Let us denote this parametrix by $\boldsymbol{\alpha}^B_{IP'PQ}$, then the medium contrast follows as

$$\mathcal{C}^{A,B}_{IP'} = \boldsymbol{\alpha}^B_{IP'PQ} [\Psi^{B'}_{PQ}] \; . \tag{7.2}$$

Substituting Eq.(6.26) into this expression leads to a direct inversion formula, with operator $\boldsymbol{\alpha}^B_{IP'PQ} \; (\boldsymbol{L}^B)^T_{PQJ}$. In the geometrical ray approximation, this operator constitutes the inverse Generalized Radon Transform (GRT), which we will denote as $\overset{\infty}{\boldsymbol{\alpha}}_{IP'PQ} \;\; (\overset{\infty}{\boldsymbol{L}})^T_{PQJ}$.

Preconditioning and image enhancing

In the method of preconditioning one introduces a linear operator, $\boldsymbol{U}^B_{IP'J}$ say, such that the composition

$$\boldsymbol{L}^B_{PIP'} \boldsymbol{U}^B_{IP'J} \quad \text{is close to the identity.} \tag{7.3}$$

The essence of preconditioning is to improve the expansion function. The original choice of expansion function, $\Phi_{IP'} = \Psi^{B'}_{P'I}$, is replaced by

$$\Phi_{PQ} = \boldsymbol{U}^B_{QPJ} [(W^T W)_{JQ'} (F^A_{Q'} - F^B_{Q'})] \; , \tag{7.4}$$

compare Eqs.(6.22) and (6.26). With this modification, the expressions for $\widehat{\mathcal{P}}$ and $\widehat{\mathcal{Q}}$, Eqs.(6.20)-(6.21), remain unchanged. The condition Eq.(7.3) allows the improvement condition that $\widehat{\mathcal{P}} \neq 0$ to hold simultaneously.

We can rewrite the procedure just described upon representing the medium update by

$$\mathcal{C}^{C,B}_{IP'} = \boldsymbol{U}^{B}_{IP'J}[H^{C,B}_{J}] \; . \tag{7.5}$$

Here, $H^{C,B}_{J}$ should be interpreted as a fictitious volume source density. Expanding this source density according to

$$H^{C,B}_{J}(\boldsymbol{x},t) = \alpha \, \Phi^{H}_{J}(\boldsymbol{x},t) \; , \tag{7.6}$$

leads to the relationship

$$\Phi_{PQ} = \boldsymbol{U}^{B}_{QPJ}[\Phi^{H}_{J}] \; . \tag{7.7}$$

This equation is then substituted into the expressions for $\widehat{\mathcal{P}}$ and $\widehat{\mathcal{Q}}$. In $\widehat{\mathcal{Q}}$ with Eq.(6.28) we thus have

$$\Psi^{\widehat{B}}_{PQ} = (\boldsymbol{L}^{B})^{T}_{PQJ}[(W^{T}W)_{JQ'}\boldsymbol{L}^{B}_{Q'IP'}[\boldsymbol{U}^{B}_{IP'J'}[\Phi^{H}_{J'}]]] \; . \tag{7.8}$$

The improvement condition is now satisfied if

$$\Phi^{H}_{J'} = (W^{T}W)_{J'Q'}(F^{A}_{Q'} - F^{B}_{Q'}) \; . \tag{7.9}$$

In choosing a preconditioner, one aims to achieve that the relative (amplitude) variations in the image matrix are more representative for the actual medium perturbation than the original image matrix. We say that the image is enhanced.

Examples of preconditioners

We will discuss two particular preconditioners; both can either be based on the geometrical ray approximation or be based on the leading-order terms of the generalized Bremmer coupling series. The idea is to let the preconditioner resemble the linearized inversion operator given by Eq.(7.2) and hence modify the expansion function to a multi-parameter-like matrix of images.

In the framework of the geometrical ray approximation, we are thus led to the choice,

$$\boldsymbol{U}^{B}_{IP'J} = \overset{\infty}{\boldsymbol{\alpha}}_{IP'PQ} \;\; (\overset{\infty}{\boldsymbol{L}})^{T}_{PQJ} \; , \tag{7.10}$$

the Generalized Radon Transform inversion. This preconditioner can be replaced by a composition of a corrector and the adjoint operator leading to the original expansion function,

$$\boldsymbol{U}^{B}_{IP'J} = \overset{\infty}{\boldsymbol{\alpha}}_{IP'PQ} \;\; (\boldsymbol{L}^{B})^{T}_{PQJ} \; , \tag{7.11}$$

cf. Eqs.(6.22) and (6.26). Esmersoy and Oristaglio [13] and Esmersoy and Miller [36] implicitly found $\overset{\infty}{\boldsymbol{\alpha}}_{IP'PQ}$ for such a preconditioner in the case of acoustic waves in two dimensions. Typically, $\overset{\infty}{\boldsymbol{\alpha}}_{IP'PQ}$ consists of an obliquity correction – which requires ray tracing to be carried out in the medium of State B to define a directional derivative along the rays associated with the field of State B', and applying it – and an amplitude correction. Sevink [37] exploited this class of preconditioners in a different way.

If the preconditioner allows the structure dual to the one in Eq.(7.11), viz.,

$$\boldsymbol{U}^{B}_{IP'J} = (\boldsymbol{L}^{B})^{T}_{IP'J'} \, V_{J'J} \ , \tag{7.12}$$

the corrector, $V_{J'J}$, can be absorbed in $(W^T W)_{JQ'}$ which brings the preconditioning into the data or observational domain. Jin *et al.* [38] investigated this corrector type, which, however, has a fundamental drawback in the sense that the parameter σ becomes the contrast-domain point in D_{con}, and hence the mismatch criterion will vary with the image point. Ignoring the obliquity part of our preconditioner Eq.(7.11) leads, in the form of Eq.(7.12), to an alternative corrector that is much simpler and has been widely used in seismic 'reverse-time' processing (see also Berkhout [39]).

We have summarized the different processors that can be built from the linearized reciprocity theorems in Table 7.1. We conjecture that the local optimization procedures based on reciprocity will yield *partial* medium reconstructions only. If we were to simplify the medium to consist of a smoothly (super-wave length scale) varying component, with perturbations (on the wave length scale) that are singular across surfaces, with (sub-wave length scale) perturbations that are random in nature, with the aid of a local optimization procedure we will typically reconstruct the singular component. The reconstruction of the unknown smoothly varying component is the subject of the following section; known fine-scale variations have to be incorporated in the pre-conditioning.

8 Combined experiments and differential semblance analysis

In reality, one commonly repeats the experiment with source distributions shifted with respect to one another. Let β parametrize the shift, then typically

$$Q^{A[\beta]}_{I}(\boldsymbol{x},t) = Q_I(\boldsymbol{x} - \boldsymbol{s}[\beta],t) \ , \quad \boldsymbol{s}[\beta] \in S_{src} \ , \ \beta \in \mathbb{R}^2 \ , \tag{8.1}$$

constitutes a family $\{A[\beta]\}$ of experiments. The vector $\boldsymbol{s}[\beta]$ will follow a curve or a surface. Associated with State $A[\beta]$ are the States $B[\beta]$ and $C[\beta]$ occurring in the optimization proced-

operator	process
$(\boldsymbol{L})^T$	imaging
$\langle(\overset{\infty}{\boldsymbol{L}})^T\ \overset{\infty}{\boldsymbol{L}}\rangle^{-1}\ (\overset{\infty}{\boldsymbol{L}})^T$	linearized GRT inversion
$\langle(\overset{\infty}{\boldsymbol{L}})^T\ \overset{\infty}{\boldsymbol{L}}\rangle^{-1}\ (\boldsymbol{L})^T$	preconditioned imaging
$\langle(\overset{\infty}{\boldsymbol{L}})^T\ \overset{\infty}{\boldsymbol{L}}\rangle^{-1}\ (\boldsymbol{L})^T\boldsymbol{L}$	resolution analysis
$\overset{\boldsymbol{r}=\boldsymbol{s}}{\boldsymbol{L}}\ \langle(\overset{\infty}{\boldsymbol{L}})^T\ \overset{\infty}{\boldsymbol{L}}\rangle^{-1}\ (\boldsymbol{L})^T$	configuration transformation

Table 7.1: Various processors (the parametrix $\boldsymbol{\alpha}$ is denoted by $\langle(\overset{\infty}{\boldsymbol{L}})^T\ \overset{\infty}{\boldsymbol{L}}\rangle^{-1}$).

ure, with

$$Q_I^{C[\beta]} = Q_I^{B[\beta]} = Q_I^{A[\beta]} \ .$$

We will integrate the different experiments,

$$A \ : \ \{A[\beta]\} \ , \quad B \ : \ \{B[\beta]\} \ , \quad C \ : \ \{C[\beta]\} \ ,$$

and introduce the mismatch between the observations in the family of States A and B based on Eq.(6.3) as

$$\epsilon^{A,B}(\sigma) = \int \mathrm{d}\beta \, \epsilon^{A[\beta],B[\beta]}(\sigma) \ . \tag{8.2}$$

Note that here D_{src} and D_{obs} become functions of β, but that D and D_{con} are experiment independent.

The local optimization condition

With the changes just introduced, we will review the local optimization procedure of Section 6. The optimization condition, Eq.(6.5), is replaced by

$$\int \mathrm{d}\beta \int_{D_{obs}[\beta]} \int_{t\in\mathbb{R}} \delta^{+}_{PJ} \, (F_P^{C[\beta]} - F_P^{B[\beta]}) \, (W^T W)_{JQ} (F_Q^{A[\beta]} - F_Q^{C[\beta]}) \, \mathrm{d}V(\boldsymbol{x}) \, \mathrm{d}t = 0 \ . \tag{8.3}$$

Like in Eq.(6.7) we will introduce States $B[\beta]'$ and $B[\beta]''$ with volume source densities

$$Q_J^{B[\beta]'} = (W^T W)_{JQ} (F_Q^{A[\beta]} - F_Q^{B[\beta]})\, \mathcal{J} \ ,$$

$$Q_J^{B[\beta]''} = (W^T W)_{JQ} (F_Q^{C[\beta]} - F_Q^{B[\beta]})\, \mathcal{J} \ ,$$

were $\mathcal{J}$ denotes a judiciously chosen Jacobian, and medium matrices

$$M_{PQ}^{B[\beta]''} = M_{PQ}^{B[\beta]'} = M_{QP}^{B[\beta]} = M_{QP}^{B} \ .$$

In the optimization condition (6.10) $\mathcal{P}$ and $\mathcal{Q}$ then have to be replaced by

$$\mathcal{P} = \int \mathrm{d}\beta \, \Big[\int_{D_{con}} \delta_{QI}^{+} \, \mathsf{R}_t(\partial_t F_P^{C[\beta]}, F_Q^{B[\beta]'})(\boldsymbol{x},0)\, \mathcal{C}_{IP}^{C,B}(\boldsymbol{x}) \, \mathrm{d}V(\boldsymbol{x})$$

$$+ \int_{\partial D} \delta_{QI}^{+} \mathcal{N}_{IP} \, \mathsf{R}_t(F_P^{C[\beta]} - F_P^{B[\beta]}, F_Q^{B[\beta]'})(\boldsymbol{x},0) \, \mathrm{d}A(\boldsymbol{x}) \Big] \ ,$$

$$\mathcal{Q} = \int \mathrm{d}\beta \, \Big[\int_{D_{con}} \delta_{QI}^{+} \, \mathsf{R}_t(\partial_t F_P^{C[\beta]}, F_Q^{B[\beta]''})(\boldsymbol{x},0)\, \mathcal{C}_{IP}^{C,B}(\boldsymbol{x}) \, \mathrm{d}V(\boldsymbol{x})$$

$$+ \int_{\partial D} \delta_{QI}^{+} \mathcal{N}_{IP} \, \mathsf{R}_t(F_P^{C[\beta]} - F_P^{B[\beta]}, F_Q^{B[\beta]''})(\boldsymbol{x},0) \, \mathrm{d}A(\boldsymbol{x}) \Big] \ , \tag{8.4}$$

compare Eq.(6.11). The medium matrices of all the different states should be independent of β.

The improvement condition and the image

The linearization yields $\mathcal{P}_{lin}$ and $\mathcal{Q}_{lin}$ as in Eq.(6.12) with modified expressions for $\Psi_{PQ}^{B'}$ and $\Psi_{PQ}^{B''}$. Equation (6.13) becomes

$$\Psi_{PQ}^{B'}(\boldsymbol{x}) = \int \mathrm{d}\beta \left[\mathcal{I}_{PQ}^{B[\beta],B[\beta]'}(\boldsymbol{x}) - \int_{\partial D} \mathcal{B}_{PQ}^{B[\beta],B[\beta]'}(\boldsymbol{x},\boldsymbol{x}') \, \mathrm{d}A(\boldsymbol{x}') \right] \ , \tag{8.5}$$

cf. Eqs.(6.14)-(6.15) and a likewise representation is found for $\Psi_{PQ}^{B''}(\boldsymbol{x})$.

To construct an approximate solution, the unknown matrix function $\mathcal{C}^{C,B}$ is again written as

$$\mathcal{C}_{IP}^{C,B}(\boldsymbol{x}) = \alpha \, \Phi_{IP}(\boldsymbol{x}) \ .$$

The family of States $\{B[\beta]''\}$ is replaced by the family $\{\widehat{B[\beta]}\}$ with

$$Q_J^{B[\beta]''} = \alpha \, Q_J^{\widehat{B[\beta]}} \ \text{ hence } \ F_Q^{B[\beta]''} = \alpha \, F_Q^{\widehat{B[\beta]}} \ , \tag{8.6}$$

which leads to the introduction of

$$\Psi^{\widehat{B}}_{PQ}(\boldsymbol{x}) = \int \mathrm{d}\beta \left[\mathcal{I}^{B[\beta],\widehat{B[\beta]}}_{PQ}(\boldsymbol{x}) - \int_{\partial D} \mathcal{B}^{B[\beta],\widehat{B[\beta]}}_{PQ}(\boldsymbol{x},\boldsymbol{x}')\,\mathrm{d}A(\boldsymbol{x}')\right] . \tag{8.7}$$

To satisfy the improvement condition, as in Eq.(6.22) we choose

$$\Phi_{IP}(\boldsymbol{x}) = \int \mathrm{d}\beta\, \Phi^{[\beta]}_{IP}(\boldsymbol{x}) = \delta^{+}_{QI}\Psi^{B'}_{PQ}(\boldsymbol{x})\,\chi_{D_{con}}(\boldsymbol{x}) \;, \tag{8.8}$$

with

$$\Phi^{[\beta]}_{IP}(\boldsymbol{x}) = \delta^{+}_{QI}\left[\mathcal{I}^{B[\beta],B[\beta]'}_{PQ}(\boldsymbol{x}) - \int_{\partial D} \mathcal{B}^{B[\beta],B[\beta]'}_{PQ}(\boldsymbol{x},\boldsymbol{x}')\,\mathrm{d}A(\boldsymbol{x}')\right]\chi_{D_{con}}(\boldsymbol{x}) \;, \tag{8.9}$$

in accordance with Eq.(8.5). The scalar α follows from Eq.(6.19) together with Eqs.(6.20)-(6.21). The degrees of freedom in the measurements $(\boldsymbol{s}[\beta],\boldsymbol{r},t) \in S_{src} \times S_{obs} \times [0,T]$ allow for a multi-parameter inversion at any $\boldsymbol{x} \in D_{con}$.

Coherency and differential semblance

The expansion function in Eq.(8.9) is the one we would find using a single experiment as in Section 6. However, the medium update from a single experiment would now become a function of $\boldsymbol{s}[\beta]$, i.e.

$$\mathcal{C}^{C,B}_{IP}(\boldsymbol{x};\beta) = \mathcal{C}^{C[\beta],B[\beta]}_{IP}(\boldsymbol{x}) = \alpha^{[\beta]}\,\Phi^{[\beta]}_{IP}(\boldsymbol{x}) \;, \tag{8.10}$$

cf. Eq.(6.16). The dependence on β reveals the degree of coherency of the reconstruction. This coherency can be measured by the *differential semblance*

$$\frac{\delta\mathcal{C}^{C[\beta],B[\beta]}_{IP}(\boldsymbol{x})}{\delta\beta} \;.$$

For a medium update to be meaningful, this differential semblance should be minimized.

To enter the constraint of maximum coherency in the optimization approach, we decompose the medium contrast $\mathcal{C}^{A,B}$ into its family members $\{\mathcal{C}^{A[\beta],B[\beta]}\}$. In accordance with our previous misfit criterion, we choose the coherency measure to be

$$\epsilon^{A,B}_{\mathcal{C}}(\sigma) = \int \mathrm{d}\beta \int_{D_{obs}[\beta]} \delta^{+}_{PJ}\,\mathsf{R}_t\Bigg([\mathbf{I}-\delta^2_\beta]^{-1/2}\,\boldsymbol{L}^{B[\beta]}_{PIP'}\left[\frac{\delta\mathcal{C}^{A[\beta],B[\beta]}_{IP'}}{\delta\beta}\right],$$

$$(W^TW)_{JQ}(\sigma)\,[\mathbf{I}-\delta^2_\beta]^{-1/2}\,\boldsymbol{L}^{B[\beta]}_{QKQ'}\left[\frac{\delta\mathcal{C}^{A[\beta],B[\beta]}_{KQ'}}{\delta\beta}\right]\Bigg)(\boldsymbol{x},0)\,\mathrm{d}V(\boldsymbol{x}) \;. \tag{8.11}$$

The resolvent of the Laplacian δ_β^2 is introduced to balance the derivatives of the medium update – inspired by the norm on the Sobolev space H^{-1} over the acquisition variables β, which is precise up to a commutator $[\boldsymbol{L}^{B[\beta]}, \delta_\beta]$.

The misfit function (cf. Eq.(6.3)) associated with the weak, linearized inversion scheme Eqs.(7.1)-(7.2) is given by

$$\epsilon_{lin}^{A[\beta],B[\beta]}(\sigma) = \int_{D_{obs}[\beta]} \delta_{PJ}^{+} \, \mathsf{R}_t((F_P^{A[\beta]} - F_P^{B[\beta]} - \boldsymbol{L}_{PIP'}^{B[\beta]}[\mathcal{C}_{IP'}^{A,B}(.;\beta)]), \tag{8.12}$$

$$(W^T W)_{JQ}(\sigma)(F_Q^{A[\beta]} - F_Q^{B[\beta]} - \boldsymbol{L}_{QKQ'}^{B[\beta]}[\mathcal{C}_{KQ'}^{A,B}(.;\beta)]))(\boldsymbol{x},0)\,\mathrm{d}V(\boldsymbol{x}) \; .$$

The differential semblance function Eq.(8.11) is then integrated with our current optimization procedure with the aid of a Lagrange multiplier, viz.,

$$\epsilon_{dso}^{A,B}(\sigma) \equiv \lambda \int \mathrm{d}\beta \, \epsilon_{lin}^{A[\beta],B[\beta]}(\sigma) + \epsilon_{\mathcal{C}}^{A,B}(\sigma) \; . \tag{8.13}$$

Optimization procedure

To make direct use of the coherency measure, we have to reformulate the optimization procedure. The use of the differential semblance misfit is improving the computational medium ('embedding') matrix in a way developed by Symes [10] and adopted to our framework here. The reformulation of the optimization problem amounts to introducing the functional

$$E(M^B;\sigma) \equiv \min_{\mathcal{C}^{A,B}} \; \epsilon_{dso}^{A,B}(\sigma) \; , \tag{8.14}$$

the minimum being the solution of the matrix integral equation

$$\int \mathrm{d}\beta \; \overset{E}{\Psi}{}_{PQ}^{B[\beta]''} = \int \mathrm{d}\beta \, \Psi_{PQ}^{B[\beta]'} \tag{8.15}$$

with

$$\overset{E}{\Psi}{}_{PQ}^{B[\beta]''} = \Big[\lambda \, (\boldsymbol{L}^{B[\beta]})_{PQJ}^T \, (W^T W)_{JQ'} \boldsymbol{L}_{Q'IP'}^{B[\beta]} \tag{8.16}$$

$$- \delta_\beta \, (\boldsymbol{L}^{B[\beta]})_{PQJ}^T \, [\mathbf{I} - \delta_\beta^2]^{-1/2} \, (W^T W)_{JQ'} [\mathbf{I} - \delta_\beta^2]^{-1/2} \, \boldsymbol{L}_{Q'IP'}^{B[\beta]} \delta_\beta \Big] \; [\mathcal{C}_{IP'}^{A,B}(.;\beta)] \; ,$$

compare Eq.(7.1). This equation follows from the substitution $\mathcal{C}^{A,B}(.;\beta) = \mathcal{C}^{A,C}(.;\beta) + \mathcal{C}^{C,B}(.;\beta)$ into $\epsilon_{dso}^{A,B}(\sigma)$; the terms linear in $\mathcal{C}^{A,C}(.;\beta)$ give the optimization condition above. Its solution can be written in terms of the parametrix of $\lambda \, (\boldsymbol{L}^{B[\beta]})_{PQJ}^T \, (W^T W)_{JQ'} \boldsymbol{L}_{Q'IP'}^{B[\beta]} - \delta_\beta \, (\boldsymbol{L}^{B[\beta]})_{PQJ}^T \, [\mathbf{I} - \delta_\beta^2]^{-1/2} \, (W^T W)_{JQ'} [\mathbf{I} - \delta_\beta^2]^{-1/2} \, \boldsymbol{L}_{Q'IP'}^{B[\beta]} \, \delta_\beta$.

Minimization of E over a class of smooth medium matrices M^B is the differential semblance optimization procedure. The more conventional semblance (or so-called migration-velocity) analysis is obtained by substituting the constraint attached to the Lagrange multiplier in the coherency measure $\epsilon_{\mathcal{C}}$. The variation $\delta\epsilon_{\mathcal{C}}$ with the medium $\mathcal{M}^B$ can then be expressed in terms of the variation of the composition of the linearized modelling and inversion operators (cf. Eq.(7.11)):

$$\delta\Big\{W\,[\mathbf{I}-\delta_\beta^2]^{-1/2}\;\boldsymbol{L}^B_{PIP'}\delta_\beta\,\boldsymbol{U}^B_{IP'J}\Big\} = W\,[\mathbf{I}-\delta_\beta^2]^{-1/2}\;\Big\{(\delta\boldsymbol{L}^B_{PIP'})\;\delta_\beta\,\boldsymbol{U}^B_{IP'J} + \boldsymbol{L}^B_{PIP'}\delta_\beta\,(\delta\boldsymbol{U}^B_{IP'J})\Big\}\;. \tag{8.17}$$

9 Discussion

We have used reciprocity to derive an optimization procedure for inverse scattering. The optimization approach has been widely used in applications. It is a reasonable approach, in particular, when the acquisition geometry with which the remote sensing is carried out is sparse, the regime where a direct inverse approach breaks down. Also, optimization schemes are attractive when a priori information is available and should be taken into account. In our paper we have critically reviewed the optimization formulation of the inverse problem and derived some refinements as well.

The iterative scheme we have proposed is of the preconditioned steepest-descent type. More sophisticated iterative schemes can be derived in a manner similar to the one discussed here. The formalism accounts for the leading non-linearity consistently with reciprocity. The resulting iterative scheme yields a series solution to the (non-linear) inverse scattering problem in case the starting value for the medium in the initial computational state is not too far from the one for the actual state. For this condition to hold, the smoothly varying component in the actual medium properties should be determined prior by other techniques; the differential semblance optimization provides a way to analyse whether the condition is satisfied and enables one to improve the smoothly varying component if necessary.

The rôle of the higher-order terms in the series or iterative solution of the inverse problem, is to compensate or 'deconvolve' for incomplete acquisition geometries and for some non-linearities such as those associated with 'interbed multiples' (for a discussion on the latter in a one-dimensional system, see Snieder [40]). These terms can improve the resolution of the final result. It is conjectured that the scale of variation of the medium should be gradually increased at each iteration by a proper adjustment of the expansion functions. The latter adjustment can be accomplished by a multi-resolution analysis.

Iterative schemes can also mimic layer stripping, a procedure arising naturally in the invariant-imbedding approach to nonlinear inversion [41]. In our formulation, we have allowed for operators that can slide time windows over the observations in accordance with pre-evaluated arrival times associated with waves scattered from image points at increasing depths: the result of a previous window is used as the initial computational medium for the subsequent window.

We kept the formulation general, in the sense that no method for computing the fields has been chosen. Approximate techniques to do so are required in practice. The well established ones are either based on dynamic ray theory ('Kirchhoff') or on the generalized Bremmer series ('wave equation'). The success of applying such techniques relies on the separability of specific non-linearities (higher-order scattering) in the inverse problem.

Acknowledgement

This research is supported in part by the Gas Research Institute and Shell Exploration and Production Co., and matching funds from the National Science Foundation Young Investigator Program, Grant No. CMS-9457268. M.V.d.H. would like to thank W.W. Symes for helpful discussions.

Appendix A. The differential operators

For *acoustic* waves, we have

$$\mathcal{D} = \begin{pmatrix} 0 & \partial_1 & \partial_2 & \partial_3 \\ \partial_1 & 0 & 0 & 0 \\ \partial_2 & 0 & 0 & 0 \\ \partial_3 & 0 & 0 & 0 \end{pmatrix}.$$

For *elastic* waves, we have

$$\mathcal{D} = \tfrac{1}{2}\mathcal{D}^{\mathrm{row/column}} + \tfrac{1}{2}\mathcal{D}^{\mathrm{diagonal}},$$

while for *electromagnetic* waves, we have

$$\mathcal{D} = \tfrac{1}{2}\mathcal{D}^{\mathrm{row/column}} - \tfrac{1}{2}\mathcal{D}^{\mathrm{diagonal}},$$

where

$$\mathcal{D}^{\mathrm{row/column}} = \left(\begin{array}{ccc} \mathbf{0} & \begin{matrix} \partial_1 & \partial_2 & \partial_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \begin{matrix} 0 & 0 & 0 \\ \partial_1 & \partial_2 & \partial_3 \\ 0 & 0 & 0 \end{matrix} & \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \partial_1 & \partial_2 & \partial_3 \end{matrix} \\ \begin{matrix} \partial_1 & 0 & 0 \\ \partial_2 & 0 & 0 \\ \partial_3 & 0 & 0 \end{matrix} & & & \\ \begin{matrix} 0 & \partial_1 & 0 \\ 0 & \partial_2 & 0 \\ 0 & \partial_3 & 0 \end{matrix} & & \mathbf{0} & \\ \begin{matrix} 0 & 0 & \partial_1 \\ 0 & 0 & \partial_2 \\ 0 & 0 & \partial_3 \end{matrix} & & & \end{array} \right) \tag{A.1}$$

and

$$\mathcal{D}^{\text{diagonal}} = \begin{pmatrix} \mathbf{0} & \begin{matrix} \partial_1 & 0 & 0 \\ 0 & \partial_1 & 0 \\ 0 & 0 & \partial_1 \end{matrix} & \begin{matrix} \partial_2 & 0 & 0 \\ 0 & \partial_2 & 0 \\ 0 & 0 & \partial_2 \end{matrix} & \begin{matrix} \partial_3 & 0 & 0 \\ 0 & \partial_3 & 0 \\ 0 & 0 & \partial_3 \end{matrix} \\ \begin{matrix} \partial_1 & 0 & 0 \\ 0 & \partial_1 & 0 \\ 0 & 0 & \partial_1 \\ \\ \partial_2 & 0 & 0 \\ 0 & \partial_2 & 0 \\ 0 & 0 & \partial_2 \\ \\ \partial_3 & 0 & 0 \\ 0 & \partial_3 & 0 \\ 0 & 0 & \partial_3 \end{matrix} & & \mathbf{0} & \end{pmatrix} . \tag{A.2}$$

Appendix B. The medium matrices

For *acoustic* waves, we have

$$M = \begin{pmatrix} \kappa & 0 & 0 & 0 \\ 0 & \rho & 0 & 0 \\ 0 & 0 & \rho & 0 \\ 0 & 0 & 0 & \rho \end{pmatrix} , \tag{B.1}$$

where ρ is the volume density of mass, and κ is the compressibility. For *elastic* waves, in terms of 3×3 submatrices, we have

$$M = \begin{pmatrix} -\rho\,\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{11} & \mathbf{S}_{12} & \mathbf{S}_{13} \\ \mathbf{0} & \mathbf{S}_{21} & \mathbf{S}_{22} & \mathbf{S}_{23} \\ \mathbf{0} & \mathbf{S}_{31} & \mathbf{S}_{32} & \mathbf{S}_{33} \end{pmatrix} , \tag{B.2}$$

where

$$(\mathbf{S}_{jl})_{ik} = s_{ijkl} \tag{B.3}$$

relates to the compliance tensor. In this notational context, it is advantageous to introduce vectorial strain rate $\dot{\mathbf{e}}_j$ and traction $\mathbf{t}_l$,

$$(\dot{\mathbf{e}}_j)_i = \tfrac{1}{2}\,(\partial_j v_i + \partial_i v_j) \ \text{ and } \ (\mathbf{t}_l)_k = \tau_{kl} \ . \tag{B.4}$$

From the symmetries of the compliance tensor, we obtain

$$\mathbf{S}_{kj} = \mathbf{S}_{jk}^T \ . \tag{B.5}$$

The stiffness tensor can be captured in a similar notation,

$$(\mathbf{C}_{jl})_{ik} = c_{ijkl} \ ; \tag{B.6}$$

the compliance and stiffness are each others inverse according to

$$\mathbf{C}_{jp}\mathbf{S}_{pk}\mathbf{t}_k = \mathbf{t}_j \ ; \quad \mathbf{C}_{jp}\mathbf{S}_{pj} = 2\,\mathbf{I} \ . \tag{B.7}$$

In the context of the *exact* contrast-source formulation Eq.(5.2) note that we can express the original compliance perturbation (Eq.(5.3)) in terms of the stiffness perturbation,

$$(\mathbf{C}_{kj}^B)^T \mathbf{S}_{kl}^{A,B} \mathbf{C}_{lm}^A = -\mathbf{C}_{jm}^{A,B} \ . \tag{B.8}$$

For *electromagnetic* waves, we have

$$M = \begin{pmatrix} \begin{matrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{21} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{31} & \epsilon_{32} & \epsilon_{33} \end{matrix} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \begin{matrix} 0 & 0 & 0 \\ 0 & \hat{\mu}_{33} & -\hat{\mu}_{32} \\ 0 & -\hat{\mu}_{23} & \hat{\mu}_{22} \end{matrix} & \begin{matrix} 0 & 0 & 0 \\ -\hat{\mu}_{33} & 0 & \hat{\mu}_{31} \\ \hat{\mu}_{23} & 0 & -\hat{\mu}_{21} \end{matrix} & \begin{matrix} 0 & 0 & 0 \\ \hat{\mu}_{32} & -\hat{\mu}_{31} & 0 \\ -\hat{\mu}_{22} & \hat{\mu}_{21} & 0 \end{matrix} \\ \mathbf{0} & \begin{matrix} 0 & -\hat{\mu}_{33} & \hat{\mu}_{32} \\ 0 & 0 & 0 \\ 0 & \hat{\mu}_{13} & -\hat{\mu}_{12} \end{matrix} & \begin{matrix} \hat{\mu}_{33} & 0 & -\hat{\mu}_{31} \\ 0 & 0 & 0 \\ -\hat{\mu}_{13} & 0 & \hat{\mu}_{11} \end{matrix} & \begin{matrix} -\hat{\mu}_{32} & \hat{\mu}_{31} & 0 \\ 0 & 0 & 0 \\ \hat{\mu}_{12} & -\hat{\mu}_{11} & 0 \end{matrix} \\ \mathbf{0} & \begin{matrix} 0 & \hat{\mu}_{23} & -\hat{\mu}_{22} \\ 0 & -\hat{\mu}_{13} & \hat{\mu}_{12} \\ 0 & 0 & 0 \end{matrix} & \begin{matrix} -\hat{\mu}_{23} & 0 & \hat{\mu}_{21} \\ \hat{\mu}_{13} & 0 & -\hat{\mu}_{11} \\ 0 & 0 & 0 \end{matrix} & \begin{matrix} \hat{\mu}_{22} & -\hat{\mu}_{21} & 0 \\ -\hat{\mu}_{12} & \hat{\mu}_{11} & 0 \\ 0 & 0 & 0 \end{matrix} \end{pmatrix} , \tag{B.9}$$

where ϵ_{ij} denotes the permittivity and $\hat{\mu}_{pq} = \frac{1}{4}\,\mu_{pq}$ with μ_{pq} denoting permeability.

Appendix C. Range of norms

In the main text of this paper, we have discussed a L^2-norm optimization scheme as it follows from the reciprocity theorem of the time-correlation type. Upon incorporating a sum of Dirac distributions in space for the observational source, we could cast the mismatch function into a ℓ^2-norm criterion. Having understood the optimization scheme in ℓ^2, it is possible to transform the scheme into a ℓ^p-norm optimization for $1 < p < 2$; the iteratively reweighted least squares algorithm provides a tool to achieve this.

In Eq.(6.3), let

$$(W^T W)_{JQ} = \chi_{[0,T]}(.)\, \frac{|S_{obs}|}{N_{obs}} \sum_{[n]=1}^{N_{obs}} S_{JQ}(\boldsymbol{r}[n], .)\, \delta_{\boldsymbol{r}[n]}\ , \quad \boldsymbol{r}[n] \in D_{obs}\ , \tag{C.1}$$

be multiplicative. Then

$$\epsilon^{A,B} = \frac{2}{p}\, \frac{|S_{obs}|}{N_{obs}} \sum_{[n]=1}^{N_{obs}} \delta^{+}_{PJ}\, \mathsf{R}_t((F_P^A - F_P^B), \chi_{[0,T]}(.)\, S_{JQ}(F_Q^A - F_Q^B))(\boldsymbol{r}[n], 0)\ . \tag{C.2}$$

By choosing

$$S_{JQ} = \begin{cases} 0 & \text{if} \quad J \neq Q \\ |F_J^A - F_J^B|^{p-2} & \text{if} \quad J = Q \end{cases} \tag{C.3}$$

the mismatch function ϵ can represent a norm in any ℓ^p; for $p = 2$ it clearly reduces to the original norm induced by reciprocity. Following Scales *et al.* [42], we use the current difference field, $F_P^A - F_P^B$, in the weighting to construct the medium of State C; we will emphasize this by the notation S_{JQ}^B. The method is referred to as *iteratively reweighted* least squares.

In practice, the weighting matrix S^B is not updated every iteration, but only evaluated after every so many steps. It is particularly useful to initiate the iterations with $p = 2$. One obvious difficulty arises near vanishing difference fields and $1 < p < 2$; for all practical purposes, a Huber taper [42] can be applied to remove this singularity.

For p less then 2, the weighting matrix tends to diminish the influence of large difference fields at isolated instants; this has the effect of eliminating the influence of outliers in the measurements and stabilizes the inversion.

References

[1] De Hoop, A.T., 1988, "Time-domain reciprocity theorems for acoustic wave fields in fluids with relaxation," *J. Acoust. Soc. Am. 84*, 1877-1882.

[2] Lions, J.L., 1968, *Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles*, Dunod-Gauthier Villars, Paris.

[3] Backus, G. and Gilbert, F., 1970, "Uniqueness in the inversion of inaccurate gross earth data," *Philos. Trans. R. Soc. London Ser. A*, 123-192.

[4] Bamberger, A., Chavent, G. and Lailly, P., 1979, "About the stability of the inverse problem in a 1-D wave equation - Application to the interpretation of seismic profiles," *J. Appl. Math. Optim. 5*, 1-47.

[5] Tarantola, A. and Valette, B., 1982, "Generalized nonlinear inverse problems solved using the least-squares criterion," *Rev. of Geoph. and Space Physics 20*, 219-232.

[6] Lailly, P., 1983, "The seismic inverse problem as a consequence of before stack migrations," in: *SIAM Conference on inverse scattering: theory and application, Edts. J.B. Bednar, R. Redner, E.A. Robinson and A.B. Weglein*, pp.206-220.

[7] Kolb, P., Collino, F. and Lailly, P., 1986, "Pre-stack inversion of a 1-D medium," *Proc. IEEE 74*, 498-508.

[8] Kennett, B.L.N. and Williamson, P.R., 1988, "Subspace methods for large-scale nonlinear inversion," in: *Mathematical Geophysics, Edts. N.J. Vlaar, G. Nolet, M.J.R. Wortel and S.A.P.L. Cloetingh*, Reidel Publishing Company, Dordrecht.

[9] Nolet, G., 1981, "Linearised inversion of teleseismic data," in: *The solution of the inverse problem in geophysical interpretation, Edt. R. Cassinis*, Plenum Press, New York.

[10] Symes, W.W., 1992, "A differential semblance criterion for inversion of multi-offset seismic reflection data," *Rice University preprint TR92-18*.

[11] Tarantola, A., 1984, "Inversion of seismic relection data in the acoustic approximation," *Geoph. 49*, 1259-1266.

[12] Miller, D., Oristaglio, M. and Beylkin, G., 1987, "A new slant on seismic imaging: migration and integral geometry," *Geoph. 52*, 943-964.

[13] Esmersoy, C. and Oristaglio, M., 1988, "Reverse-time wave-field extrapolation, imaging, and inversion," *Geoph. 53*, 920-931.

[14] Claerbout, J.F., 1971, "Toward a unified theory of reflector mapping," *Geoph. 36*, 467-481.

[15] Claerbout, J.F., 1992, *Earth sounding analysis: processing versus inversion*, Blackwell Scientific Publications.

[16] Berkhout, A.J., 1982, *Imaging acoustic energy by wave field extrapolation*, 2nd edition, Elsevier, Amsterdam.

[17] Wapenaar, C.P.A. and Berkhout, A.J., 1989, *Elastic wave field extrapolation*, Elsevier, Amsterdam.

[18] Stolt, R.H. and Weglein, A.B., 1985, "Migration and inversion of seismic data," *Geoph. 50*, 2458-2472.

[19] Bleistein, N., 1987, "On the imaging of reflectors in the earth," *Geoph. 52*, 931-942.

[20] Parker, R.L., 1977, "Understanding inverse theory," *Ann. Rev. Earth Planet. Sci. 5*, 35-64.

[21] Aki, K. and Richards, P.G., 1980, *Quantitative seismology*, Freeman and Co., San Francisco.

[22] De Hoop, A.T., 1995, *Handbook of radiation and scattering of waves*, Academic Press, London.

[23] Woodhouse, J.H., 1974, "Surface waves in a laterally varying layered structure," *Geophys. J. R. Astr. Soc. 37*, 461-490.

[24] De Hoop, M.V., 1992, *Directional decomposition of transient acoustic wave fields*, Delft University Press, Delft.

[25] De Hoop, A.T. and De Hoop, M.V., 1996, "Acoustic, elastodynamic and electromagnetic wavefield computation - a structured approach based on reciprocity," in: *Proceedings* of the symposium 'Large-scale structures in acoustics and electromagnetics', National Academic Press, Washington D.C., pp.72-88.

[26] Chapman C.H. and Coates, R.T., 1994, "Generalized Born scattering in anisotropic media," *Wave Motion 14*, 309-341.

[27] De Hoop, A.T., 1991, "Convergence criterion for the time-domain iterative Born approximation to scattering by an inhomogeneous, dispersive object," *J. Opt. Soc. Am. A/8*, 1256-1260.

[28] Rakesh, 1988, "A linearised inverse problem for the wave equation," *Comm. in Part. Diff. Eqs. 13*, 573-601.

[29] Hagedoorn, J.G., 1954, "A process of seismic reflection interpretation," *Geoph. Prosp. 2*, 85-127.

[30] Beylkin, G., 1985, "Imaging of discontinuities in the inverse scattering problem by inversion of the causal generalized Radon transform," *J. Math. Phys. 26*, 99-108.

[31] Berkhout, A.J. and Rietveld, W.E.A., 1995, "Prestack migration in terms of double dynamic focusing," *65th Ann. Intern. Mtg. Soc. Expl. Geophys.*, Expanded Abstracts, pp.1228-1231.

[32] Thorbecke, J., 1997,*Common Focus Point technology*, PhD Thesis, Delft University of Technology.

[33] Schultz, P.S. and Sherwood, J.W.C., 1980, "Depth migration before stack," *Geoph. 45*, 376-393.

[34] Berryhill, J.R., 1984, "Wave equation datuming before stack," *Geoph. 49*, 2064-2066.

[35] Bunks, C., Saleck, F.M., Zaleski, S. and Chavent, G., 1995, "Multiscale seismic waveform inversion," *Geoph. 60*, 1457-1473.

[36] Esmersoy, C. and Miller, D., 1989, "Backprojection versus backpropagation in multidimensional linearized inversion," *Geoph. 54*, 921-926.

[37] Sevink, G.J.A., 1996, *Asymptotic seismic inversion*, PhD Thesis, Delft University of Technology, Delft.

[38] Jin, S., Madariaga, R., Virieux, J. and Lambaré, G., 1992, "Two-dimensional asymptotic iterative elastic inversion," *Geophys. J. Int. 108*, 575-588.

[39] Berkhout, A.J., 1982, *Seismic migration, Developments in solid earth geophysics 14A*, Elsevier, Amsterdam.

[40] Snieder, R., 1990, "A perturbative analysis of nonlinear inversion," *Geoph. J. Int. 101*, 545-556.

[41] Weston, V.H., 1992, "Invariant imbedding and wave splitting in $\mathbb{R}^3$: II. The Green function approach to inverse scattering," *Inv. Problems 8*, 919-947.

[42] Scales, J.A., Gersztenkorn, A. and Treitel, S., 1988, "Fast l_p solution of large, sparse, linear systems: application to seismic travel time tomography," *J. Comp. Phys. 75*, 314-335.