Tutorial on two-way wave equation operators for acoustic, isotropic, constant-density media

Guillaume Barnier and Ali Almomin

ABSTRACT

This paper is a tutorial on linearized two-way wave equation modeling and inversion operators. We provide a detailed derivation for the special case of an acoustic, isotropic, constant-density medium. We analyze the Born, tomographic, and WEMVA forward modeling operators, their adjoints, and we extend the analysis to the subsurface offset domain.

INTRODUCTION

We present a detailed derivation for the main two-way wave equation linearized operators used in seismic imaging: Born, tomography, and Wave Equation Migration Velocity Analysis (WEMVA). This work adapts that of Almomin (2013) and is intended to serve as an educational tool for new students coming to SEP. We will consider the specific case of an acoustic, isotropic, constant-density medium.

Our goal is to obtain the best possible estimate of the seismic velocity model of the Earth's subsurface. The three operators that we derive are convenient to achieve this goal. The Born operator and its adjoint capture the dynamic effects responsible for seismic reflections (the high wavenumber content of the velocity model). The tomographic and WEMVA operators and their adjoints capture the kinematic effects (e.g., transmitted and diving waves), which are controlled by the low wavenumber content of the velocity model.

The first section is intended to remind the reader of some general background on wave theory. We then derive the Born modeling operator and its adjoint, referred to as the Reverse Time Migration (RTM) operator. In the last two sections, we will treat the tomographic and WEMVA operators and their respective adjoints.

Throughout this paper, we use roman fonts to refer to functions, italic fonts to refer to functions evaluated at a given point, and bold fonts to refer to vectors (e.g., f, $f(x)$, and $\mathbf{f}$). Moreover, we use lowercase to refer to functions in the time domain, and uppercase for their Discrete Fourier Transform (DFT).

WAVE THEORY

Definitions

- m is a function representing the seismic velocity model of the subsurface. For each point in the subsurface, it associates a velocity value:

$$\begin{aligned} \mathrm{m} : \Omega &\mapsto \mathbb{R} \\ \mathbf{x} &\mapsto m(\mathbf{x}), \end{aligned} \tag{1}$$

  where $m(\mathbf{x})$ is expressed in m/s. $\Omega \subset \mathbb{R}^3$ is the area of study, and is an open, bounded and regular set. It is assumed that m does not vary with time.

- p is the pressure field function. For a given location in the subsurface $\mathbf{x}$, at a given time t, and for a given velocity function m, it represents the pressure value

$$\begin{aligned} \mathrm{p} : \Omega \times \mathbb{R} \times \mathcal{F}(\Omega, \mathbb{R}) &\mapsto \mathbb{R} \\ (\mathbf{x}, t; \mathrm{m}) &\mapsto p(\mathbf{x}, t; \mathrm{m}), \end{aligned} \tag{2}$$

  where $p(\mathbf{x}, t; \mathrm{m})$ is expressed in Pa, and $\mathcal{F}(\Omega, \mathbb{R})$ is the set of functions mapping Ω to $\mathbb{R}$ (assumed to be infinitely differentiable on Ω).

- s is a function representing a seismic source:

$$\begin{aligned} \mathrm{s} : \Omega \times \mathbb{R} &\mapsto \mathbb{R} \\ (\mathbf{x}, t) &\mapsto s(\mathbf{x}, t), \end{aligned} \tag{3}$$

  where $s(\mathbf{x}, t)$ is expressed in Pa/m^2.

- ρ is a function representing the medium volume mass density and is defined by

$$\begin{aligned} \rho : \Omega &\mapsto \mathbb{R} \\ \mathbf{x} &\mapsto \rho(\mathbf{x}), \end{aligned} \tag{4}$$

  where $\rho(\mathbf{x})$ is expressed in kg/m^3. In the following, we assume ρ to be constant over the study area.

- $\forall \mathbf{x} \in \Omega, \forall t \in \mathbb{R}$, the acoustic, isotropic constant-density wave equation satisfied by the causal function p is given by

$$\begin{aligned} \frac{1}{m^2(\mathbf{x})} \frac{\partial^2 p(\mathbf{x}, t; \mathrm{m})}{\partial t^2} - \nabla^2 p(\mathbf{x}, t; \mathrm{m}) &= s(\mathbf{x}, t) \\ p(\mathbf{x}, t; \mathrm{m}) &\equiv 0, \ \forall t \leq 0. \end{aligned} \tag{5}$$

Continuous solution of the wave equation

Green's function

Given a source located at $\mathbf{x}'$ and initiated at time t', the causal Green's function, g, for an acoustic, isotropic constant-density medium is the solution to equation 5, where s is a point source function, and an impulse in time. That is, $\forall \mathbf{x} \in \Omega,\ \forall t \in \mathbb{R}, \forall t' \in \mathbb{R}$, g satisfies

$$\begin{aligned} \frac{1}{m^2(\mathbf{x})}\,\frac{\partial^2 g(\mathbf{x},t,\mathbf{x}',t\,';\mathrm{m})}{\partial t^2} - \nabla^2 g(\mathbf{x},t,\mathbf{x}',t';\mathrm{m}) &= \delta(\mathbf{x}-\mathbf{x}')\delta(t-t') \\ g(\mathbf{x},t,\mathbf{x}',t';\mathrm{m}) &\equiv 0,\ \forall t \leq t\,', \end{aligned} \tag{6}$$

where

- $\delta(\mathbf{x}-\mathbf{x}')$ is in m^{-3} (for 3D),
- $\delta(t-t')$ is in s^{-1}, and
- $g(\mathbf{x},t,\mathbf{x}',t';\mathrm{m})$ is in $\mathrm{m}^{-1}\,\mathrm{s}^{-1}$ (and not Pa).

Spatial reciprocity implies that

$$g(\mathbf{x},t,\mathbf{x}',t\,';\mathrm{m}) = g(\mathbf{x}',t,\mathbf{x},t\,';\mathrm{m}). \tag{7}$$

Assuming the properties of the medium are invariant with time, we can write

$$g(\mathbf{x},t,\mathbf{x}',t\,';\mathrm{m}) = \mathrm{g}(\mathbf{x},t-t\,',\mathbf{x}',0;\mathrm{m}). \tag{8}$$

$\forall \mathbf{x} \in \Omega, \forall t \in \mathbb{R}$, the solution of equation 5 can be written as

$$p(\mathbf{x},t;\mathrm{m}) = \int\limits_{\mathbf{x}'\in\Omega}\ \int\limits_{t'=-\infty}^{+\infty} g(\mathbf{x},t,\mathbf{x}',t\,';\mathrm{m})\ s(\mathbf{x}',t')\ \mathrm{d}t'\ \mathrm{d}\mathbf{x}'. \tag{9}$$

Using equation 8, we may write

$$\int\limits_{t'=-\infty}^{+\infty} g(\mathbf{x},t,\mathbf{x}',t\,';\mathrm{m})\ s(\mathbf{x}',t\,')\ \mathrm{d}t\,' = g(\mathbf{x},t,\mathbf{x}',0;\mathrm{m}) * s(\mathbf{x}',t), \tag{10}$$

where * denotes convolution in time. Equation 9 thereby simplifies to

$$p(\mathbf{x},t;\mathrm{m}) = \int\limits_{\mathbf{x}'\in\Omega} g(\mathbf{x},t,\mathbf{x}',0;\mathrm{m}) * s(\mathbf{x}',t)\ \mathrm{d}\mathbf{x}'. \tag{11}$$

If the seismic source function s is concentrated at one point in space $\mathbf{x}_s$, and has a time signature $f(t)$, then $s(\mathbf{x},t) = f(t)\ \delta(\mathbf{x}-\mathbf{x}_s)$, where $f(t)$ is expressed in Pa m, and $\delta(\mathbf{x}-\mathbf{x}_s)$ is expressed in m^{-3}. In that case, equation 11 simplifies further to

$$p(\mathbf{x},t;\mathrm{m}) = g(\mathbf{x},t,\mathbf{x}_s;\mathrm{m}) * f(t),\ \forall \mathbf{x}\in\Omega, \forall t\in\mathbb{R}. \tag{12}$$

Numerical solution of the wave equation

Discretization in time and frequency

We discretize all functions/signals in time. Hence $\forall n \in \{1; N\}$,

$$p(\mathbf{x},n;\mathrm{m}) \quad = \quad p(\mathbf{x},t_n;\mathrm{m}), \tag{13}$$

where dt is the time sampling rate in s/samples, and $t_n = (n-1)\,dt$. Here, it is assumed that $\exists N \in \mathbb{N}$ such that $n \notin \{1; N\} \Rightarrow p(\mathbf{x},n;\mathrm{m}) = 0$.

By taking the Discrete Fourier Transform (DFT) of each side of equation 11, we obtain that $\forall k \in \{1; N\}$,

$$P(\mathbf{x},\omega_k;\mathrm{m}) = \int\limits_{\mathbf{x}'\in\Omega} G(\mathbf{x},\omega_k,\mathbf{x}';\mathrm{m})\ S(\mathbf{x}',\omega_k)\ \mathrm{d}\mathbf{x}', \tag{14}$$

where $\omega_k = (k-1)d\omega = (k-1)\dfrac{2\pi}{N}$. Note that $P(\mathbf{x},\omega_k;\mathrm{m}) \in \mathbb{C}$. Moreover,

- $P(\mathbf{x},\omega_k;\mathrm{m})$ is expressed in Pa s,
- $F(\omega_k)$ is expressed in Pa m s, and
- $G(\mathbf{x},\omega_k,\mathbf{x}_s;\mathrm{m})$ is expressed in m^{-1}.

Discretization in space

We discretize the area of study into a regularly sampled grid. $\forall n \in \{1; N\}$, $\forall j \in \{1; M\}$, equation 11 becomes

$$p(\mathbf{x}, n; \mathbf{m}) \approx \sum_{j=1}^{M} g(\mathbf{x}, n, \mathbf{x}_j; \mathrm{m}) * s(\mathbf{x}_j, n)\ \Delta\mathbf{x}, \tag{15}$$

where $\Delta\mathbf{x}$ is the constant finite difference grid cell volume, expressed in m^3. Since $\Delta\mathbf{x}$ acts as a constant scaling factor, we will not write it explicitly in the remaining of our derivation. However, this coefficient is important to ensure consistency in the units. The bold font used for the model function $\mathbf{m}$ indicates that we have spatially discretized the study area. M is the number of grid points in the study area (discretized model size). Therefore, we can define a model vector $\mathbf{m} \in \mathbb{R}^M$, whose components are the values of function m (defined in mapping 1), evaluated at each grid point $\mathbf{x}_i$

$$\mathbf{m} = \begin{pmatrix} m(\mathbf{x}_1) \\ \vdots \\ m(\mathbf{x}_M) \end{pmatrix} = \begin{pmatrix} m_1 \\ \vdots \\ m_M \end{pmatrix}. \tag{16}$$

We can also express equation 15 in the frequency domain. $\forall i \in \{1; M\}$, $\forall k \in \{1; N\}$,

$$P(\mathbf{x}_i, \omega_k; \mathbf{m}) \ \approx \ \sum_{j=1}^{M} G(\mathbf{x}_i, \omega_k, \mathbf{x}_j; \mathbf{m})\ S(\mathbf{x}_j, \omega_k). \tag{17}$$

If the seismic acquisition source is concentrated at a point $\mathbf{x}_s$ in space, its DFT is expressed by $S(\mathbf{x}_i, \omega_k) = F(\omega_k)\ \delta(\mathbf{x}_i - \mathbf{x}_s)$, where F is the DFT of the source's time signature. In that case, equation 17 simplifies further to

$$P(\mathbf{x}_i, \omega_k; \mathbf{m}) = G(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{m})\ F(\omega_k). \tag{18}$$

Seismic data

Seismic data are a set of discrete measurements in time and space *received* at some locations $\mathbf{x}_i$ (e.g., near the surface of the Earth). For one seismic source located at $\mathbf{x}_s$, we defined a data vector $\mathbf{D} \in \mathbb{C}^{N_d \times N}$ (N_d is the number of receiver locations)

$$\mathbf{D} = \begin{pmatrix} D_{11} \\ \vdots \\ D_{N_d N} \end{pmatrix}. \tag{19}$$

Each component D_{ik} of $\mathbf{D}$ can be expressed by

$$D_{ik} = P(\mathbf{x} = \mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{m}), \tag{20}$$

where $i \in \{1; N_d\},\ k \in \{1; N\}$.

Wave equation and Green's function in the frequency domain

By discretizing in space and taking the DFT of equation 6, the Green's function G satisfies the Helmholtz equation. $\forall i \in \{1; M\}, \forall k \in \{1; N\}$,

$$\left[\frac{\omega_k^2}{m(\mathbf{x}_i)^2} + \nabla^2\right] G(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{m}) = -\delta(\mathbf{x}_i - \mathbf{x}_s). \tag{21}$$

The units of equation 21 are consistent, as each side is expressed in m^{-3}.

LINEARIZATION WITH RESPECT TO REFLECTIVITY

Nonlinear mapping

One way to think about the wave equation is as a function f that maps the set of model parameters (velocity value at each grid point) to the recorded seismic data

$$\begin{aligned} \mathrm{f} : \mathbb{R}^M &\mapsto \mathbb{R}^{N_d} \\ \mathbf{m} &\mapsto \mathbf{f}(\mathbf{m}), \end{aligned} \tag{22}$$

where $\mathbf{d} = \mathbf{f}(\mathbf{m})$ is the seismic data in the time domain. Equivalently, in the frequency domain,

$$\begin{aligned} \mathrm{F} : \mathbb{R}^M &\mapsto \mathbb{C}^{N_d} \\ \mathbf{m} &\mapsto \mathbf{F}(\mathbf{m}), \end{aligned} \tag{23}$$

where $\mathbf{D} = \mathbf{F}(\mathbf{m})$ is the seismic data in the frequency domain. Clearly, f and F (to be distinguished from the source signatures mentioned previously) are not linear functions with respect to $\mathbf{m}$. However, they are both linear operators with respect to the *source function*, keeping the velocity model and all other variables unchanged. Let us represent the wave equation operator by $\mathcal{L}$ such that

$$\begin{aligned} \mathcal{L} : \mathcal{F}(\Omega, \mathbb{R}) &\mapsto \mathcal{F}(\Omega, \mathbb{R}) \\ \mathrm{s} &\mapsto \mathcal{L}(\mathrm{s}). \end{aligned} \tag{24}$$

Here, $\mathrm{s} \in \mathcal{F}(\Omega, \mathbb{R})$ is a source function. One can easily verify from equation 5 that, $\forall \alpha \in \mathbb{R}, \forall\ (\mathrm{s}_i, \mathrm{s}_j) \in (\mathcal{F}(\Omega, \mathbb{R}))^2$

$$\begin{aligned} \mathcal{L}(\mathrm{s}_i + \mathrm{s}_j) &= \mathcal{L}(\mathrm{s}_i) + \mathcal{L}(\mathrm{s}_j) \\ \mathcal{L}(\alpha\, \mathrm{s}_i) &= \alpha \mathcal{L}(\mathrm{s}_i). \end{aligned} \tag{25}$$

The next step is to linearize F with respect to $\mathbf{m}$. In the following, we present our derivation in the frequency domain. First, we assume that we can decompose the model vector $\mathbf{m} \in \mathbb{R}^M$ as the sum of two vectors

$$\mathbf{m} = \mathbf{b} + \mathbf{r}, \tag{26}$$

where

$$\mathbf{b} = \begin{pmatrix} b(\mathbf{x}_1) \\ \vdots \\ b(\mathbf{x}_M) \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_M \end{pmatrix}, \tag{27}$$

and

$$\mathbf{r} = \begin{pmatrix} r(\mathbf{x}_1) \\ \vdots \\ r(\mathbf{x}_M) \end{pmatrix} = \begin{pmatrix} r_1 \\ \vdots \\ r_M \end{pmatrix}. \tag{28}$$

$\mathbf{b} \in \mathbb{R}^M$ is referred to as the *background* model, and contains the low wavenumber content of the velocity model. $\mathbf{r} \in \mathbb{R}^M$ is referred to as the *reflectivity* model, and contains the high wavenumber content of the velocity model. We also assume that the magnitude of the reflectivity is much smaller than the one of the background, $\|\mathbf{r}\| \ll \|\mathbf{b}\|$, where $\|.\|$ denotes any norm on $\mathbb{R}^M$. In the following, we will treat those parameters as two separate variables with same units ($\mathrm{m\,s^{-1}}$).

Let us consider the function

$$\begin{aligned} \tilde{\mathrm{F}} : \mathbb{R}^M &\mapsto \mathbb{C}^{N_d} \\ \mathbf{r} &\mapsto \tilde{\mathbf{F}}(\mathbf{r}), \end{aligned} \tag{29}$$

which is the restriction of F to the high wavenumber part of the velocity model, while keeping the background velocity model unchanged. We perform a multivariate Taylor expansion of $\tilde{\mathrm{F}}$ around a reflectivity model $\mathbf{r}_0$ (referred to as the background reflectivity) while keeping the background model $\mathbf{b}$ unchanged

$$\tilde{\mathbf{F}}(\mathbf{r}) = \tilde{\mathbf{F}}(\mathbf{r}_0) + \left.\frac{\partial \tilde{\mathbf{F}}(\mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{r}_0} \Delta\mathbf{r} + \mathcal{O}(\|\Delta\mathbf{r}\|^2). \tag{30}$$

The second term of the right side of equation 30 is the Jacobian matrix $\mathcal{B}$ of $\tilde{\mathbf{F}}$ evaluated at $\mathbf{r} = \mathbf{r}_0$, applied to the reflectivity perturbation vector $\Delta\mathbf{r}$. It can be expressed by

$$\mathcal{B}(\mathbf{r}_0) = \left.\frac{\partial \tilde{\mathbf{F}}(\mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{r}_0} = \begin{pmatrix} \frac{\partial \tilde{F}_1}{\partial r_1}\Big|_{\mathbf{r}_0} & \cdots & \frac{\partial \tilde{F}_1}{\partial r_M}\Big|_{\mathbf{r}_0} \\ \vdots & \ddots & \vdots \\ \frac{\partial \tilde{F}_{N_d}}{\partial r_1}\Big|_{\mathbf{r}_0} & \cdots & \frac{\partial \tilde{F}_{N_d}}{\partial r_M}\Big|_{\mathbf{r}_0} \end{pmatrix}, \tag{31}$$

where $\mathcal{B}(\mathbf{r}_0) \in M_{N_d,M}(\mathbb{C})$ ($M_{p,q}(\mathbb{C})$ refers to the set of matrices with p rows, q columns, and complex coefficients).

Born operator

The Born operator is a linear operator that relates a perturbation in the model reflectivity to a perturbation in the data (Almomin, 2013), while keeping the background unchanged.

Expressing the entries of the Jacobian

We consider a reflectivity perturbation $\Delta\mathbf{r}$ vector such that

- $\mathbf{r} = \mathbf{r}_0 + \Delta\mathbf{r}$, and
- $\|\Delta\mathbf{r}\| \ll \|\mathbf{r}_0\|$.

Thus,

$$\begin{aligned} \Delta\mathbf{D} &= \mathbf{D}(\mathbf{r}) - \mathbf{D}(\mathbf{r}_0) \\ &\approx \mathcal{B}(\mathbf{r}_0)\, \Delta\mathbf{r}. \end{aligned} \tag{32}$$

The Born operator is a linear application $\mathcal{B}(\mathbf{r}_0) : \mathbb{C}^M \mapsto \mathbb{C}^{N_d}$, which can be represented in a matrix form by $\mathcal{B}(\mathbf{r}_0)$. It is the Jacobian of $\tilde{\mathbf{F}}$ taken at $\mathbf{r}_0$, whose entries are all independent of the perturbation $\Delta\mathbf{r}$.

Let us define, $\forall i \in \{1; N_d\}, \forall k \in \{1; N\}$,

$$D_{ik}(\mathbf{b},\mathbf{r}) = P(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r}) = F(\omega_k)\, G(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r}), \tag{33}$$

where $D_{ik}(\mathbf{b},\mathbf{r}) \in \mathbb{C}$. Therefore, we can rewrite equation 32 with the following form:

$$\begin{aligned} \Delta D(\mathbf{x}_i, \mathbf{x}_s, \omega_k, ; \mathbf{b}, \mathbf{r}) &= P(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r}) - P(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r}_0) \\ &\approx \sum_{j=1}^{M} \frac{\partial P}{\partial r_j}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r})\Big|_{\mathbf{r}=\mathbf{r}_0} \Delta r(\mathbf{x}_j) \\ &\approx F(\omega_k) \sum_{j=1}^{M} \frac{\partial G}{\partial r_j}(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r})\Big|_{\mathbf{r}=\mathbf{r}_0} \Delta r(\mathbf{x}_j). \end{aligned} \tag{34}$$

The left side of equation 34 is the data perturbation caused by the perturbation of the reflectivity model, for one fixed frequency, one fixed receiver location, one fixed source, and a fixed background velocity model.

To find an expression of $\dfrac{\partial G}{\partial r_j}$, we write two versions of equation 21: one with the unperturbed velocity model $\mathbf{m}_0 = \mathbf{b} + \mathbf{r}_0$, satisfied by $G(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r}_0)$ (equation 35), and the other with the perturbed velocity model $\mathbf{m} = \mathbf{m}_0 + \Delta\mathbf{r}$, satisfied by $G(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r})$. That gives us, $\forall i \in \{1; M\}, \forall k \in \{1; N\}$,

$$\left[\frac{\omega_k^2}{\big(m_0(\mathbf{x}_i) + \Delta r(\mathbf{x}_i)\big)^2} + \nabla^2\right] G(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r}) = -\delta(\mathbf{x}_i - \mathbf{x}_s), \tag{35}$$

and

$$\left[\frac{\omega_k^2}{m_0(\mathbf{x}_i)^2} + \nabla^2\right] G(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r}_0) \quad = \quad -\delta(\mathbf{x}_i - \mathbf{x}_s). \tag{36}$$

Since $\forall i \in \{1; M\}$, $|\Delta r(\mathbf{x}_i)| \ll |m_0(\mathbf{x}_i)|$, we can make the following approximation:

$$\frac{1}{\big(m_0(\mathbf{x}_i) + \Delta r(\mathbf{x}_i)\big)^2} \approx \frac{1}{m_0(\mathbf{x}_i)^2}\left(1 - \frac{2\Delta r(\mathbf{x}_i)}{m_0(\mathbf{x}_i)}\right). \tag{37}$$

By expanding equations 35, 36, and 37, we can show that the difference $\Delta G(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r}) = G(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r}) - G(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r}_0)$ satisfies a similar equation as in equation 21, but for a different source function. $\forall i \in \{1; M\}, \forall k \in \{1; N\}$,

$$\begin{aligned}\left[\frac{\omega_k^2}{m_0(\mathbf{x}_i)^2} + \nabla^2\right] \Delta G(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r}) = \\ 2\,\omega_k^2\, \frac{\Delta r(\mathbf{x}_i)}{m_0(\mathbf{x}_i)^3}\, \Big(G(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r}_0) + \Delta G(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r})\Big).\end{aligned} \tag{38}$$

Under that form, this wave equation is not linear with respect to the source (mapping 24). However, we can further simplify it by noticing in equation 38 that the term proportional to $\Delta r(\mathbf{x}_i)\Delta G$ is a second order differential element, and can thus be neglected with respect to the remaining terms (Born approximation). Hence,

$$|\Delta r(\mathbf{x}_i)\Delta G| \ll |\Delta r(\mathbf{x}_i)\, G_0|, \tag{39}$$

and

$$\left|\frac{2\,\omega_k^2}{m_0(\mathbf{x}_i)^3}\Delta r(\mathbf{x}_i)\Delta G\right| \ll \left|\left[\frac{\omega_k^2}{m_0(\mathbf{x}_i)^2} + \nabla^2\right]\Delta G\right|. \tag{40}$$

Therefore, equation 38 simplifies to

$$\left[\frac{\omega_k^2}{m_0(\mathbf{x}_i)^2} + \nabla^2\right] \Delta G \approx S_{\mathrm{sec}}(\mathbf{x}_i, \omega_k), \tag{41}$$

with

$$S_{\mathrm{sec}}(\mathbf{x}_i,\omega_k) = 2\,\omega_k^2\,\frac{\Delta r(\mathbf{x}_i)}{m_0(\mathbf{x}_i)^3}\;G(\mathbf{x}_i,\omega_k,\mathbf{x}_s;\mathbf{b},\mathbf{r}_0). \tag{42}$$

ΔG satisfies the wave equation with a secondary source function $\mathrm{S}_{\mathrm{sec}}$, *independent* of ΔG, and proportional to the reflectivity perturbation at a given point in the subsurface. Using equation 17, we can solve for ΔG. $\forall i \in \{1;M\}, \forall k \in \{1;N\}$,

$$\begin{aligned}\Delta G(\mathbf{x}_i,\omega_k;\mathbf{b},\mathbf{r}) &= \sum_{j=1}^{M} G(\mathbf{x}_i,\omega_k,\mathbf{x}_j;\mathbf{b},\mathbf{r}_0)\;S_{\mathrm{sec}}(\mathbf{x}_j,\omega_k) \\ &= \sum_{j=1}^{M}\frac{2\,\omega_k^2}{m_0(\mathbf{x}_j)^3}\,G(\mathbf{x}_i,\omega_k,\mathbf{x}_j;\mathbf{b},\mathbf{r}_0)\,G(\mathbf{x}_j,\omega_k,\mathbf{x}_s;\mathbf{b},\mathbf{r}_0)\;\Delta r(\mathbf{x}_j).\end{aligned} \tag{43}$$

We check for the consistency of the units in equation 43, keeping in mind that we did not explicitly write the grid cell volume term (expressed in m^3) on the right side,

- $\Delta G(\mathbf{x}_i,\omega_k;\mathbf{b},\mathbf{r})$ is expressed in m^{-1},
- $G(\mathbf{x}_i,\omega_k,\mathbf{x}_j;\mathbf{b},\mathbf{r}_0)$ is in m^{-1}, and
- $S_{\mathrm{sec}}(\mathbf{x}_i,\omega_k)$ is in m^{-3}.

The Taylor expansion of ΔG, expressed by

$$\begin{aligned}\Delta G(\mathbf{x}_i,\omega_k,\mathbf{x}_s;\mathbf{b},\mathbf{r}) &= G(\mathbf{x}_i,\omega_k,\mathbf{x}_s;\mathbf{b},\mathbf{r}) - G(\mathbf{x}_i,\omega_k,\mathbf{x}_s;\mathbf{b},\mathbf{r}_0) \\ &\approx \sum_{j=1}^{M}\frac{\partial G}{\partial r_j}(\mathbf{x}_i,\omega_k;\mathbf{x}_s;\mathbf{b},\mathbf{r})\bigg|_{\mathbf{r}=\mathbf{r}_0}\;\Delta r(\mathbf{x}_j),\end{aligned} \tag{44}$$

enables us to identify $\dfrac{\partial G}{\partial r_j}$, where

$$\frac{\partial G}{\partial r_j}(\mathbf{x}_i,\omega_k;\mathbf{x}_s;\mathbf{b},\mathbf{r})\bigg|_{\mathbf{r}=\mathbf{r}_0} = \frac{2\,\omega_k^2}{m_0(\mathbf{x}_j)^3}\,G(\mathbf{x}_i,\omega_k,\mathbf{x}_j;\mathbf{b},\mathbf{r}_0)\,G(\mathbf{x}_j,\omega_k,\mathbf{x}_s;\mathbf{b},\mathbf{r}_0). \tag{45}$$

Finally, we can express the perturbation of the data with respect to the perturbation of the reflectivity. $\forall i \in \{1;N_d\}, \forall k \in \{1;N\}$,

$$\begin{aligned}\Delta D_{ik} &= \Delta D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r}) \\ &= \sum_{j=1}^{M} F(\omega_k) \frac{2\,\omega_k^2}{m_0(\mathbf{x}_j)^3} G(\mathbf{x}_i, \omega_k, \mathbf{x}_j; \mathbf{b}, \mathbf{r}_0) \Delta r(\mathbf{x}_j)\, G(\mathbf{x}_j, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r}_0).\end{aligned} \tag{46}$$

Equation 46 can be written in a matrix form

$$\Delta \mathbf{D} = \mathcal{B}(\mathbf{r}_0) \Delta \mathbf{r}, \tag{47}$$

and

$$\Delta \mathbf{D} = \begin{pmatrix} \Delta D(\mathbf{x}_1, \mathbf{x}_s, \omega_1; \mathbf{b}, \mathbf{r}) \\ \Delta D(\mathbf{x}_2, \mathbf{x}_s, \omega_1; \mathbf{b}, \mathbf{r}) \\ \vdots \\ \Delta D(\mathbf{x}_{N_d}, \mathbf{x}_s, \omega_1; \mathbf{b}, \mathbf{r}) \\ \vdots \\ \Delta D(\mathbf{x}_1, \mathbf{x}_s, \omega_N; \mathbf{b}, \mathbf{r}) \\ \vdots \\ \Delta D(\mathbf{x}_{N_d}, \mathbf{x}_s, \omega_N; \mathbf{b}, \mathbf{r}) \end{pmatrix} = \begin{pmatrix} b_{11}^1 & \cdots & b_{11}^M \\ b_{21}^1 & \cdots & b_{21}^M \\ \vdots & \vdots & \vdots \\ b_{N_d 1}^1 & \cdots & b_{N_d 1}^M \\ \vdots & \vdots & \vdots \\ b_{1N}^1 & \cdots & b_{1N}^M \\ \vdots & \vdots & \vdots \\ b_{N_d N}^1 & \cdots & b_{N_d N}^M \end{pmatrix} \begin{pmatrix} \Delta r(\mathbf{x}_1) \\ \Delta r(\mathbf{x}_2) \\ \vdots \\ \Delta r(\mathbf{x}_M) \end{pmatrix}, \tag{48}$$

where $\mathcal{B}(\mathbf{r}_0)$ is the Born modeling operator, linearized around $\mathbf{m}_0 = \mathbf{b} + \mathbf{r}_0$. Moreover,

- $\Delta \mathbf{D} \in \mathbb{C}^{N_d \times N}$,
- $\mathcal{B}(\mathbf{r}_0) \in M_{N_d, M}(\mathbb{C})$,
- $b_{ik}^{\,j} \in \mathbb{C}$, and
- $\Delta \mathbf{r} \in \mathbb{R}^M$.

Each entry $b_{ik}^{\,j}$ of $\mathcal{B}(\mathbf{r}_0)$ is independent of the reflectivity perturbation $\Delta \mathbf{r}$, and can be expressed $\forall i \in \{1; N_d\}, \forall k \in \{1; N\}, \forall j \in \{1; M\}$ by

$$b_{ik}^{\,j} = F(\omega_k) \frac{2\,\omega_k^2}{m_0(\mathbf{x}_j)^3} G(\mathbf{x}_i, \omega_k, \mathbf{x}_j; \mathbf{b}, \mathbf{r}_0)\, G(\mathbf{x}_j, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r}_0). \tag{49}$$

Extension to subsurface offset domain

We extend the reflectivity part of the model to the subsurface offset domain (Almomin, 2013), and define an extended reflectivity function $\tilde{\mathrm{r}}$ by

$$\begin{aligned} \tilde{\mathrm{r}} : \Omega \times \mathcal{H} &\mapsto \mathbb{R} \\ \mathbf{x} \times \mathbf{h} &\mapsto \tilde{r}(\mathbf{x}, \mathbf{h}), \end{aligned} \tag{50}$$

where both Ω and $\mathcal{H}$ are subsets of $\mathbb{R}^3$. We discretize the domain on which $\tilde{\mathrm{r}}$ is defined as in equation 28, and we obtain the reflectivity vector $\tilde{\mathbf{r}} \in \mathbb{R}^{M \times N_h}$ extended to the subsurface offset domain, where $N_h = 2\,h + 1$. Moreover, $\forall p \in \{-h; h\}$, $\mathbf{h}_{\text{-}p} = -\mathbf{h}_p$, and

$$\tilde{\mathbf{r}} = \begin{pmatrix} \tilde{r}(\mathbf{x}_1, \mathbf{h}_{\text{-}h}) \\ \vdots \\ \tilde{r}(\mathbf{x}_M, \mathbf{h}_{\text{-}h}) \\ \tilde{r}(\mathbf{x}_1, \mathbf{h}_{\text{-}h+1}) \\ \vdots \\ \tilde{r}(\mathbf{x}_M, \mathbf{h}_h) \end{pmatrix} = \begin{pmatrix} \tilde{r}(\mathbf{x}_1, -\mathbf{h}_h) \\ \vdots \\ \tilde{r}(\mathbf{x}_M, -\mathbf{h}_h) \\ \tilde{r}(\mathbf{x}_1, -\mathbf{h}_{h\text{-}1}) \\ \vdots \\ \tilde{r}(\mathbf{x}_M, \mathbf{h}_h) \end{pmatrix} \tag{51}$$

Equation 46 can then be modified to take into account the subsurface offset dimension (Almomin, 2013). $\forall i \in \{1; N_d\}, \forall k \in \{1; N\}$, we have

$$\begin{aligned} \Delta D_{ik} &= \Delta D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}) \\ &= \sum_{j=1}^{M} \sum_{p=\text{-}h}^{h} F(\omega_k) \frac{2\,\omega_k^2}{m_0(\mathbf{x}_j)^3}\, G(\mathbf{x}_i, \omega_k, \mathbf{x}_j + \mathbf{h}_p; \mathbf{b}, \tilde{\mathbf{r}}_0) \Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p)\, G(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s; \mathbf{b}, \tilde{\mathbf{r}}_0). \end{aligned} \tag{52}$$

We can display equation 52 in the matrix form

$$\Delta \mathbf{D} = \tilde{\mathcal{B}}(\tilde{\mathbf{r}}_0)\; \Delta\tilde{\mathbf{r}}, \tag{53}$$

and

$$\Delta\mathbf{D} = \begin{pmatrix} \Delta D(\mathbf{x}_1, \mathbf{x}_s, \omega_1; \mathbf{b}, \tilde{\mathbf{r}}) \\ \Delta D(\mathbf{x}_2, \mathbf{x}_s, \omega_1; \mathbf{b}, \tilde{\mathbf{r}}) \\ \vdots \\ \Delta D(\mathbf{x}_{N_d}, \mathbf{x}_s, \omega_1; \mathbf{b}, \tilde{\mathbf{r}}) \\ \Delta D(\mathbf{x}_1, \mathbf{x}_s, \omega_2; \mathbf{b}, \tilde{\mathbf{r}}) \\ \vdots \\ \Delta D(\mathbf{x}_1, \mathbf{x}_s, \omega_N; \mathbf{b}, \tilde{\mathbf{r}}) \\ \vdots \\ \Delta D(\mathbf{x}_{N_d}, \mathbf{x}_s, \omega_N; \mathbf{b}, \tilde{\mathbf{r}}) \end{pmatrix} = \begin{pmatrix} \tilde{b}_{11}^{1(\text{-}h)} & \cdots & \tilde{b}_{11}^{M(\text{-}h)} & \tilde{b}_{11}^{1(\text{-}h+1)} & \cdots & \tilde{b}_{11}^{Mh} \\ \tilde{b}_{21}^{1(\text{-}h)} & \cdots & \tilde{b}_{21}^{M(\text{-}h)} & \tilde{b}_{21}^{1(\text{-}h+1)} & \cdots & \tilde{b}_{21}^{Mh} \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ \tilde{b}_{N_d1}^{1(\text{-}h)} & \cdots & \tilde{b}_{N_d1}^{M(\text{-}h)} & \tilde{b}_{N_d1}^{1(\text{-}h+1)} & \cdots & \tilde{b}_{N_d1}^{Mh} \\ \tilde{b}_{12}^{1(\text{-}h)} & \cdots & \tilde{b}_{12}^{M(\text{-}h)} & \tilde{b}_{12}^{1(\text{-}h+1)} & \cdots & \tilde{b}_{12}^{Mh} \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ \tilde{b}_{1N}^{1(\text{-}h)} & \cdots & \tilde{b}_{1N}^{M(\text{-}h)} & \tilde{b}_{1N}^{1(\text{-}h+1)} & \cdots & \tilde{b}_{1N}^{Mh} \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ \tilde{b}_{N_dN}^{1(\text{-}h)} & \cdots & \tilde{b}_{N_dN}^{M(\text{-}h)} & \tilde{b}_{N_dN}^{1(\text{-}h+1)} & \cdots & \tilde{b}_{N_dN}^{Mh} \end{pmatrix} \begin{pmatrix} \Delta\tilde{r}(\mathbf{x}_1, \mathbf{h}_{\text{-}h}) \\ \vdots \\ \Delta\tilde{r}(\mathbf{x}_M, \mathbf{h}_{\text{-}h}) \\ \Delta\tilde{r}(\mathbf{x}_1, \mathbf{h}_{\text{-}h+1}) \\ \vdots \\ \Delta\tilde{r}(\mathbf{x}_M, \mathbf{h}_h) \end{pmatrix}, \quad (54)$$

where $\tilde{\mathcal{B}}(\tilde{\mathbf{r}}_0)$ is the Born operator in the extended domain. Moreover,

- $\Delta\mathbf{D} \in \mathbb{C}^{N_d \times N}$,
- $\tilde{\mathcal{B}}(\tilde{\mathbf{r}}_0) \in M_{N_d \times N, M \times N_h}(\mathbb{C})$,
- $\tilde{b}_{ik}^{\,jp} \in \mathbb{C}$, and
- $\Delta\tilde{\mathbf{r}} \in \mathbb{R}^{M \times N_h}$.

We can explicitly write the expressions for the entries of $\tilde{\mathcal{B}}(\tilde{\mathbf{r}}_0)$. $\forall i \in \{1; N_d\}, \forall k \in \{1; N\}, \forall j \in \{1; M\}, \forall p \in \{-h; h\}$, each entry $\tilde{b}_{ik}^{\,jp}$ is given by

$$\tilde{b}_{ik}^{\,jp} = F(\omega_k) \frac{2\,\omega_k^2}{m_0(\mathbf{x}_j)^{\,3}} G(\mathbf{x}_i, \omega_k, \mathbf{x}_j + \mathbf{h}_p; \mathbf{b}, \tilde{\mathbf{r}}_0)\, G(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s; \mathbf{b}, \tilde{\mathbf{r}}_0). \quad (55)$$

We have obtained the expression for the Born modeling operator in the extended subsurface offset domain (equations 53, 54, and 55), which relates a perturbation in the reflectivity model to a perturbation in the data, keeping all other parameters unchanged, and assuming a known $\mathbf{m}_0$.

Time domain expression of the Born operator

To get a better physical understanding of the Born operator, it is convenient to express our previous results in the time domain. By taking the inverse DFT of equation 43, we can express the data perturbation as the convolution in time between the Green's function and a secondary source (caused by the reflectivity perturbation). $\forall i \in \{1; N_d\}, \forall n \in \{1; N\}, \forall j \in \{1; M\}, \forall p \in \{-h; h\}$, we have

$$\Delta d(\mathbf{x}_i, n; \mathbf{b}, \mathbf{r}) \;=\; \sum_{p=\text{-}h}^{h} \sum_{j=1}^{M} g(\mathbf{x}_i, n, \mathbf{x}_j + \mathbf{h}_p, 0; \mathbf{b}, \mathbf{r}_0) * s_{\text{sec}}(\mathbf{x}_j, \mathbf{h}_p, n). \tag{56}$$

The secondary source can be further expressed by

$$\begin{aligned} s_{\text{sec}}(\mathbf{x}_j, \mathbf{h}_p, n) &= \text{DFT}^{-1}\big[S_{\text{sec}}(\mathbf{x}_j, \mathbf{h}_p, \omega_k)\big] \\ &= \text{DFT}^{-1}\left[\frac{2F(\omega_k)\;\omega_k^2}{m_0(\mathbf{x}_j)^3}\Delta\tilde{\mathrm{r}}(\mathbf{x}_j, \mathbf{h}_p)\;\; G(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s; \mathbf{b}, \mathbf{r}_0)\right] \\ &= p_{\text{src}}(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_s, n; \mathbf{b}, \mathbf{r}_0)\;\Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p), \end{aligned} \tag{57}$$

where,

$$p_{\text{src}}(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_s, n; \mathbf{b}, \mathbf{r}_0) = \frac{-2}{m_0(\mathbf{x}_j)^3}\;\ddot{f}(n) * g(\mathbf{x}_j - \mathbf{h}_p, n, \mathbf{x}_s, 0; \mathbf{b}, \mathbf{r}_0). \tag{58}$$

Function $\ddot{\mathrm{f}}$ is the discrete second order time derivative of the original source time signature. Function p_{src}, referred to as source wavefield, is a scaled version of the time convolution between the second time derivative of the source signature and the Green's function computed with the known velocity model $\mathbf{m}_0 = \mathbf{b} + \mathbf{r}_0$. $\forall i \in \{1; N_d\}, \forall n \in \{1; N\}$, the data perturbation can therefore be expressed by

$$\begin{aligned} &\Delta d(\mathbf{x}_i, n; \mathbf{b}, \mathbf{r}) = \\ &\quad \sum_{p=\text{-}h}^{h} \sum_{j=1}^{M} g(\mathbf{x}_i, n, \mathbf{x}_j + \mathbf{h}_p, 0; \mathbf{b}, \mathbf{r}_0) * \big[p_{\text{src}}(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_s, n; \mathbf{b}, \mathbf{r}_0)\;\Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p)\big]. \end{aligned} \tag{59}$$

Physical interpretation of Born modeling operator

For simplicity, we focus our interpretation on the specific case of zero-subsurface offset. Let us consider the scenario where we have

- a single seismic source located at a point $\mathbf{x}_s$ at the surface,
- a subsurface location $\mathbf{x}_j$ such that $\Delta r(\mathbf{x}_j) \neq 0$, and
- a single recording location $\mathbf{x}_i$ at the surface, where we would like to compute the data perturbation $\Delta d(\mathbf{x}_i, n; \mathbf{b}, \mathbf{r})$.

The source wavefield $\mathrm{p}_{\mathrm{src}}$ generated at location $\mathbf{x}_s$ is propagated into the subsurface with a known velocity model $\mathbf{m}_0$. The secondary source $\mathrm{s}_{\mathrm{sec}}$ generated at $\mathbf{x}_j$, is the product of the source wavefield $\mathrm{p}_{\mathrm{src}}$ with the reflectivity perturbation $\Delta r(\mathbf{x}_j)$ (equation 57). $\mathrm{s}_{\mathrm{sec}}$ is nonzero if and only if the reflectivity perturbation is nonzero. In equation 59, the convolution between the secondary source and the Green's function indicates that a secondary wavefield (referred to as the scattered wavefield $\mathrm{p}_{\mathrm{scat}}$) is generated from the secondary source. Therefore, the contribution of the reflectivity perturbation $\Delta r(\mathbf{x}_j)$ to the data perturbation $\Delta d(\mathbf{x}_i, n; \mathbf{b}, \mathbf{r})$ is obtained by extracting the values of $\mathrm{p}_{\mathrm{scat}}$ at location $\mathbf{x}_i$. This process is illustrated in Figure 1. Finally, in order to capture the contributions from all the reflectivity perturbations in the subsurface to an observation point $\mathbf{x}_i$, we sum over all subsurface points $\mathbf{x}_j$ (equation 59).

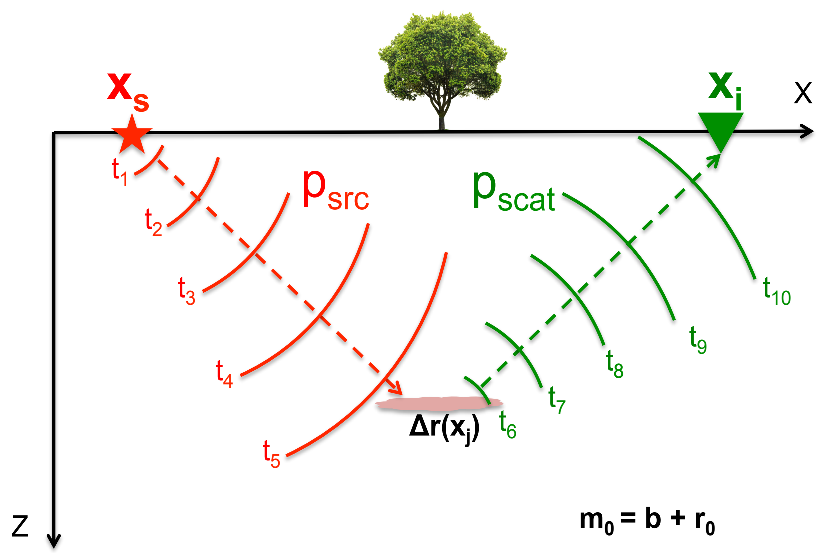


Figure 1: Schematic diagram of the Born modeling operator for one scattering point, and a known background velocity model $\mathbf{m}_0$. The source wavefield $\mathrm{p}_{\mathrm{src}}$ (red) interacts with the reflectivity perturbation $\Delta r(\mathbf{x}_j)$ (pink), and creates a scattered wavefield $\mathrm{p}_{\mathrm{scat}}$ (green). [**NR**]

RTM

The RTM operator is defined as the adjoint of the Born modeling operator. From equation 54, we can obtain the adjoint $\tilde{\mathcal{R}}(\tilde{\mathbf{r}}_0)$ of operator $\tilde{\mathcal{B}}(\tilde{\mathbf{r}}_0)$ by taking its conjugate transpose matrix, $\tilde{\mathcal{R}}(\tilde{\mathbf{r}}_0) = \tilde{\mathcal{B}}(\tilde{\mathbf{r}}_0)^*$, which satisfies

$$\Delta\tilde{\mathbf{r}} = \tilde{\mathcal{R}}(\tilde{\mathbf{r}}_0)\,\Delta\mathbf{D}, \tag{60}$$

and

$$\begin{pmatrix} \Delta\tilde{r}(\mathbf{x}_1, \mathbf{h}_{\text{-}h}) \\ \vdots \\ \Delta\tilde{r}(\mathbf{x}_M, \mathbf{h}_{\text{-}h}) \\ \Delta\tilde{r}(\mathbf{x}_1, \mathbf{h}_{\text{-}h+1}) \\ \vdots \\ \Delta\tilde{r}(\mathbf{x}_M, \mathbf{h}_{h}) \end{pmatrix} = \begin{pmatrix} \tilde{r}_{11}^{1(\text{-}h)} & \cdots & \tilde{r}_{N_d1}^{1(\text{-}h)} & \cdots & \tilde{r}_{1N}^{1(\text{-}h)} & \cdots & \tilde{r}_{N_dN}^{1(\text{-}h)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \tilde{r}_{11}^{M(\text{-}h)} & \cdots & \tilde{r}_{N_d1}^{M(\text{-}h)} & \cdots & \tilde{r}_{1N}^{M(\text{-}h)} & \cdots & \tilde{r}_{N_dN}^{M(\text{-}h)} \\ \tilde{r}_{11}^{1(\text{-}h+1)} & \cdots & \tilde{r}_{N_d1}^{1(\text{-}h+1)} & \cdots & \tilde{r}_{1N}^{1(\text{-}h+1)} & \cdots & \tilde{r}_{N_dN}^{1(\text{-}h+1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \tilde{r}_{11}^{Mh} & \cdots & \tilde{r}_{N_d1}^{Mh} & \cdots & \tilde{r}_{1N}^{Mh} & \cdots & \tilde{r}_{N_dN}^{Mh} \end{pmatrix} \begin{pmatrix} \Delta D(\mathbf{x}_1, \omega_1) \\ \vdots \\ \Delta D(\mathbf{x}_{N_d}, \omega_1) \\ \vdots \\ \Delta D(\mathbf{x}_1, \omega_N) \\ \vdots \\ \Delta D(\mathbf{x}_{N_d}, \omega_N) \end{pmatrix}, \quad (61)$$

where

- $\Delta\tilde{\mathbf{r}} \in \mathbb{C}^{M\times N_h}$,
- $\tilde{\mathcal{R}}(\tilde{\mathbf{r}}_0) \in M_{M\times N_h, N_d\times N}(\mathbb{C})$,
- $\tilde{r}\,^{jp}_{ik} = (\tilde{b}\,^{jp}_{ik})^* \in \mathbb{C}$, and
- $\Delta\mathbf{D} \in \mathbb{R}^{M\times N_h}$.

Moreover, each entry $\tilde{r}\,^{jp}_{ik}$ of $\tilde{\mathcal{R}}(\tilde{\mathbf{r}}_0)$ is given by

$$\begin{aligned} \tilde{r}\,^{jp}_{ik} &= (\tilde{b}\,^{jp}_{ik})\,^* & (62) \\ &= F^*(\omega_k)\frac{2\,\omega_k^2}{m_0(\mathbf{x}_j)\,^3}\,G^*(\mathbf{x}_i, \omega_k, \mathbf{x}_j + \mathbf{h}_p; \mathbf{b}, \tilde{\mathbf{r}}_0)\,G^*(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s; \mathbf{b}, \tilde{\mathbf{r}}_0). \end{aligned}$$

Therefore, one row of equation 54 is given by

$$\begin{aligned} &\Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p) = & (63) \\ &\quad \sum_{i=1}^{N_d}\sum_{k=1}^{N} \frac{2F^*(\omega_k)\,\omega_k^2}{m_0(\mathbf{x}_j)^3}\,G^*(\mathbf{x}_i, \omega_k, \mathbf{x}_j + \mathbf{h}_p; \mathbf{b}, \tilde{\mathbf{r}}_0)\Delta D(\mathbf{x}_i, \omega_k)\,G^*(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s; \mathbf{b}, \tilde{\mathbf{r}}_0). \end{aligned}$$

We have obtained the expression for the RTM operator in the extended subsurface offset domain (equations 61, 62, and 63). It is the adjoint of the Born modeling operator, and it relates a perturbation in the data to a perturbation in the reflectivity model, while keeping other parameters unchanged.

Time domain expression of RTM operator

In a similar fashion as for the Born operator, we express the results obtained for the RTM operator in the time domain in order to get a better understanding of its physical meaning. $\forall j \in \{1; M\}, \forall p \in \{-h; h\}$, equation 63 can be rewritten as

$$\Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p) = \sum_{k=1}^{N} P^*_{\text{src}}(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_s, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}_0)\ P_{\text{rec}}(\mathbf{x}_j, \mathbf{h}_p, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}_0), \tag{64}$$

where:

- $P_{\text{src}}(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_s, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}_0) = \text{DFT}\big[p_{\text{src}}(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_s, n; \mathbf{b}, \tilde{\mathbf{r}}_0)\big]$, and
- $P_{\text{rec}}(\mathbf{x}_j, \mathbf{h}_p, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}_0) = \sum_{i=1}^{N_d} G^*(\mathbf{x}_i, \omega_k, \mathbf{x}_j + \mathbf{h}_p; \mathbf{b}, \tilde{\mathbf{r}}_0)\Delta D(\mathbf{x}_i, \omega_k)$.

P_{src} is the DFT of the source wavefield p_{src} previously defined. P_{rec} is the DFT of the *receiver wavefield*, defined as the time convolution between the data perturbation and the anti-causal Green's function g_-:

$$p_{\text{rec}}(\mathbf{x}_j, \mathbf{h}_p, n; \mathbf{b}, \tilde{\mathbf{r}}_0) = \sum_{i=1}^{N_d} g_-(\mathbf{x}_j + \mathbf{h}_p, n, \mathbf{x}_i, 0; \mathbf{b}, \tilde{\mathbf{r}}_0) * \Delta d(\mathbf{x}_i, n; \mathbf{b}, \mathbf{r}). \tag{65}$$

Therefore the receiver wavefield is the progagation *backward in time* of the data perturbation. Using equation 121 (appendix), we can show that

$$\begin{aligned}
&\big[p_{\text{src}} \otimes p_{\text{rec}}\big](\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_s, 0; \mathbf{b}, \tilde{\mathbf{r}}_0) \\
&= \sum_{k=1}^{N} P^*_{\text{src}}(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_s, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}_0)\ P_{\text{rec}}(\mathbf{x}_j, \mathbf{h}_p, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}_0) \\
&= \sum_{n=1}^{N} p_{\text{src}}(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_s, n; \mathbf{b}, \tilde{\mathbf{r}}_0)\ p_{\text{rec}}(\mathbf{x}_j, \mathbf{h}_p, n; \mathbf{b}, \tilde{\mathbf{r}}_0),
\end{aligned} \tag{66}$$

where the left side of equation 66 is the zero-lag time cross-correlation of the source wavefield p_{src} with the receiver wavefield p_{rec} at various subsurface locations. For any subsurface point $\mathbf{x}_j$ and for any subsurface offset $\mathbf{h}_p$, we can now express the reflectivity perturbation by

$$\Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p) = \big[p_{\text{src}} \otimes p_{\text{rec}}\big](\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_s, 0; \mathbf{b}, \tilde{\mathbf{r}}_0), \tag{67}$$

where $\Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p) \in \mathbb{R}$.

Physical interpretation of RTM operator

For simplicity, let us physically interpret the specific case of zero-subsurface offset, and let us consider

- a single source located at a point $\mathbf{x}_s$ at the surface, and
- a single recording location $\mathbf{x}_i$ at the surface where we have a data perturbation $\Delta D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r})$.

We apply the RTM operator in order to recover the location(s) of the reflectivity perturbation(s) that caused the data perturbation observed at the surface. The source wavefield $\mathrm{p}_{\mathrm{src}}$ generated at location $\mathbf{x}_s$ is propagated forward in time with the known velocity model $\mathbf{m}_0$. The receiver wavefield $\mathrm{p}_{\mathrm{rec}}$ generated by the data perturbation $\Delta D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r})$ at $\mathbf{x}_i$ is propagated backward in time with the known velocity $\mathbf{m}_0$. For each point in the subsurface, the reflectivity perturbation value is equal to the output of the cross-correlation of the two wavefields at zero-time lag. If there are locations such that both wavefields coincide at the same time, the cross-correlation output will be non-zero (assuming there is not only destructive interferences), and a reflectivity perturbation will be generated at these locations. This process is illustrated in Figure 2. Finally, to account for all the contributions coming from other potential observation and source locations, the reflectivity perturbations computed for each source/observation pair are summed, and a reflectivity perturbation map is generated. It is referred to as an *image*.

It is common in RTM to linearize the wave equation around a background reflectivity $\mathbf{r}_0 = \mathbf{0}$, which means that the model $\mathbf{m}_0$ used to compute the source wavefield, scattered wavefield, and receiver wavefield contains only low wavenumber components. In other words, a smooth velocity model is used to obtain the RTM image. If that is the case,

$$\Delta D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r}) \quad = \quad D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \Delta\mathbf{r}) - D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{0}), \tag{68}$$

where $D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{0})$ is the modeled data using a smooth background velocity model, which contains only direct arrivals and diving waves (no reflections). Hence, the data perturbation $\Delta D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r})$ used to compute the receiver wavefield will consist of the recorded data $D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \Delta\mathbf{r})$ from which we have removed the direct arrivals and diving waves. Though confusing, it is common to call $\Delta D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r})$ the "data."

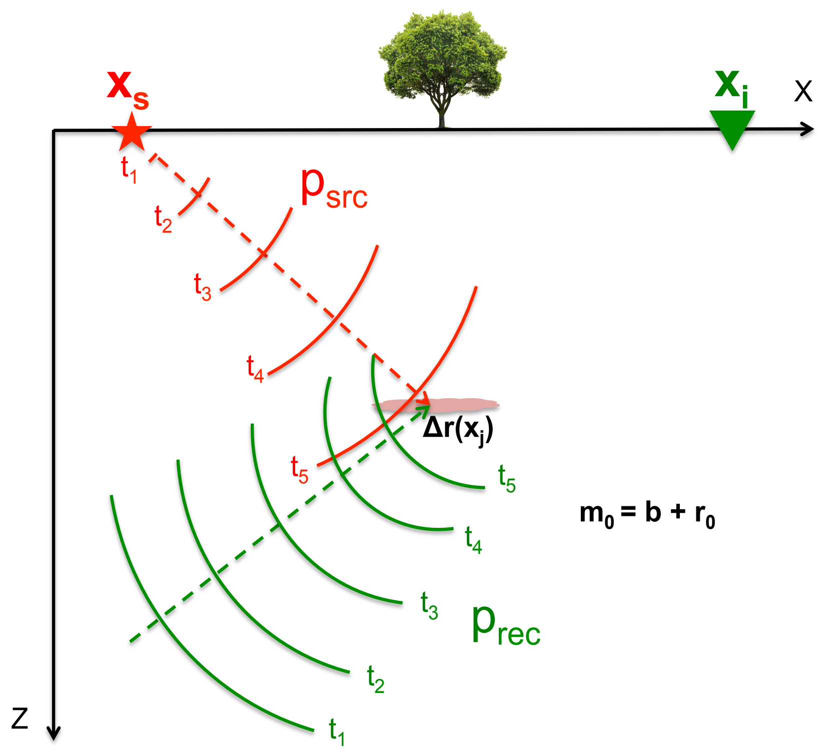


Figure 2: Schematic diagram of the RTM operator for one source located at $\mathbf{x}_s$, and one receiver located at $\mathbf{x}_i$. For all points in the subsurface, the source wavefield $\mathrm{p_{src}}$ (red) is cross-correlated at zero-time lag with the receiver wavefield $\mathrm{p_{rec}}$ (green). The receiver wavefield is propagated backward in time from the receiver location $\mathbf{x}_i$. A nonzero reflectivity perturbation will be generated at the subsurface points where the two wavefields coincide in time and space. [**NR**]

LINEARIZATION WITH RESPECT TO THE BACKGROUND

In previous sections, we linearized the two-way wave equation with respect to the high wavenumber part of the velocity model (reflectivity), while keeping the lower wavenumber part (background) unchanged. In this section, we show how the tomographic and WEMVA operators are obtained by linearizing the Born and RTM operators with respect to the background velocity model.

Tomographic operator

The forward tomographic operator, as defined in Almomin (2013), is a linear operator that relates a perturbation in the background velocity model $\Delta\mathbf{b}$ (such that $\mathbf{b} = \mathbf{b}_0 + \Delta\mathbf{b}$) to a perturbation in the perturbation of the data $\Delta(\Delta\mathbf{D})$, while keeping the other parameters unchanged. We can symbolically write it as

$$\Delta(\Delta\mathbf{D}) = \mathbf{T}(\mathbf{b}_0, \tilde{\mathbf{r}})\Delta\mathbf{b}. \tag{69}$$

During this linearization process, we will assume that the reflectivity model $\tilde{\mathbf{r}} = \tilde{\mathbf{r}}_0 + \Delta\tilde{\mathbf{r}}$ (i.e., the image) is unchanged, and part of the operator.

Linearization of the Born operator

In order to obtain the expression for the tomographic operator, we perturb the background velocity model $\mathbf{b}$, and we express the perturbation in the perturbation of the data. We first define $\tilde{D}(\mathbf{x}_i, \omega_k, \mathbf{x}_s; \mathbf{b}, \tilde{\mathbf{r}})$ by

$$\begin{aligned}\tilde{D}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}) &= \Delta D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}) \\ &= D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}) - D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}_{\mathbf{0}}).\end{aligned} \tag{70}$$

We perform a multivariate first-order Taylor expansion of the data perturbation $\tilde{D}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \tilde{\mathbf{r}})$ around a known model $\mathbf{b}_0$, such that $\mathbf{b} = \mathbf{b}_0 + \Delta\mathbf{b}$. $\mathbf{b}_0$ is referred to as the *background background model*, and $\Delta\mathbf{b}$ is the *background perturbation*. Moreover, throughout this process, $\tilde{\mathbf{r}} = \tilde{\mathbf{r}}_0 + \Delta\tilde{\mathbf{r}}$ is assumed to be known and is kept unchanged. Assuming a small background perturbation $\Delta\mathbf{b}$, we can write

$$\begin{aligned}\Delta\tilde{D}_{ik} = \Delta\tilde{D}(\mathbf{x}_i, \mathbf{x}_s\omega_k, \mathbf{b}, \tilde{\mathbf{r}}) &= \tilde{D}(\mathbf{x}_i, \mathbf{x}_s, \omega_k, ; \mathbf{b}, \tilde{\mathbf{r}}) - \tilde{D}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}_0, \tilde{\mathbf{r}}) \\ &\approx \sum_{q=1}^{M} \frac{\partial \tilde{D}}{\partial b_q}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \tilde{\mathbf{r}})\bigg|_{\mathbf{b}=\mathbf{b}_0} \Delta b(\mathbf{x}_q).\end{aligned} \tag{71}$$

Using equation 52, we have

$$\begin{aligned} \frac{\partial \tilde{D}}{\partial b_q}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \tilde{\mathbf{r}})\bigg|_{\mathbf{b}=\mathbf{b}_0} &= \frac{\partial \Delta D}{\partial b_q}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r})\bigg|_{\mathbf{b}=\mathbf{b}_0} \\ &= \sum_{j=1}^{M}\sum_{p=-h}^{h} \frac{\partial \beta_T}{\partial b_q}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{h}_p, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}_0)\bigg|_{\mathbf{b}=\mathbf{b}_0} \alpha_T(\mathbf{x}_j, \mathbf{h}_p, \omega_k), \end{aligned} \tag{72}$$

where

- $\alpha_T(\mathbf{x}_j, \mathbf{h}_p, \omega_k) = F(\omega_k)\dfrac{2\,\omega_k^2}{m_0(\mathbf{x}_j)^3}\Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p)$, and
- $\beta_T(\mathbf{x}_i, \mathbf{x}_j, \mathbf{h}_p, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}_0) = G(\mathbf{x}_i, \omega_k, \mathbf{x}_j + \mathbf{h}_p; \mathbf{b}, \tilde{\mathbf{r}}_0)\, G(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s; \mathbf{b}, \tilde{\mathbf{r}}_0)$.

Therefore

$$\begin{aligned} \frac{\partial \beta_T}{\partial b_q}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{h}_p, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}_0)\bigg|_{\mathbf{b}_0} &= \frac{\partial G}{\partial b_q}(\mathbf{x}_i, \omega_k, \mathbf{x}_j + \mathbf{h}_p; \mathbf{b}, \tilde{\mathbf{r}}_0)\bigg|_{\mathbf{b}_0} G(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s; \mathbf{b}_0, \tilde{\mathbf{r}}_0) \\ &+ G(\mathbf{x}_i, \omega_k, \mathbf{x}_j + \mathbf{h}_p; \mathbf{b}_0, \tilde{\mathbf{r}}_0)\frac{\partial G}{\partial b_q}(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s; \mathbf{b}, \tilde{\mathbf{r}}_0)\bigg|_{\mathbf{b}_0}. \end{aligned}$$

We can obtain $\dfrac{\partial G}{\partial b_q}(\mathbf{x}_i, \omega_k, \mathbf{x}_j + \mathbf{h}_p; \mathbf{b}, \tilde{\mathbf{r}}_0)\bigg|_{\mathbf{b}_0}$ and $\dfrac{\partial G}{\partial b_q}(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s; \mathbf{b}, \tilde{\mathbf{r}}_0)\bigg|_{\mathbf{b}_0}$ the same way as in previous section (equation 45) by setting $\mathbf{m}_0 = \mathbf{b}_0 + \mathbf{r}_0$. Hence, $\forall i \in \{1; M\}, \forall k \in \{1; N\}, \forall j \in \{1; M\}, \forall q \in \{1; M\}, \forall p \in \{-h; h\}$,

$$\begin{aligned} &\frac{\partial G}{\partial b_q}(\mathbf{x}_i, \omega_k, \mathbf{x}_j + \mathbf{h}_p; \mathbf{b}, \tilde{\mathbf{r}}_0)\bigg|_{\mathbf{b}_0} = \\ &\quad \frac{2\,\omega_k^2}{m_0(\mathbf{x}_q)^3} G(\mathbf{x}_i, \omega_k, \mathbf{x}_q; \mathbf{b}_0, \tilde{\mathbf{r}}_0)\, G(\mathbf{x}_q, \omega_k, \mathbf{x}_j + \mathbf{h}_p; \mathbf{b}_0, \tilde{\mathbf{r}}_0), \end{aligned} \tag{73}$$

and

$$\begin{aligned} &\frac{\partial G}{\partial b_q}(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s; \mathbf{b}, \tilde{\mathbf{r}}_0)\bigg|_{\mathbf{b}_0} = \\ &\quad \frac{2\,\omega_k^2}{m_0(\mathbf{x}_q)^3} G(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_q; \mathbf{b}_0, \tilde{\mathbf{r}}_0)\, G(\mathbf{x}_q, \omega_k, \mathbf{x}_s; \mathbf{b}_0, \tilde{\mathbf{r}}_0). \end{aligned} \tag{74}$$

$\forall i \in \{1; N_d\}, \forall k \in \{1; N\}$, equation 71 now becomes

$$\Delta\tilde{D}_{ik} = \Delta\tilde{D}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}) = \sum_{q=1}^{M} T_{ik}^{\,q}\ \Delta b(\mathbf{x}_q), \tag{75}$$

with $T_{ik}^{\,q} \in \mathbb{C}$, and

$$T_{ik}^{\,q} = L_1(\mathbf{x}_i, \mathbf{x}_q, \omega_k) + L_2(\mathbf{x}_i, \mathbf{x}_q, \omega_k), \tag{76}$$

where

$$L_1(\mathbf{x}_i, \mathbf{x}_q, \omega_k) = \sum_{j=1}^{M} \sum_{p=-h}^{h} \gamma_{\,jq}^{\,k} G_0(\mathbf{x}_i, \omega_k, \mathbf{x}_q)\, G_0(\mathbf{x}_q, \omega_k, \mathbf{x}_j + \mathbf{h}_p) \Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p)\, G_0(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s), \tag{77}$$

and

$$L_2(\mathbf{x}_i, \mathbf{x}_q, \omega_k) = \sum_{j=1}^{M} \sum_{p=-h}^{h} \gamma_{\,jq}^{\,k} G_0(\mathbf{x}_i, \omega_k, \mathbf{x}_j + \mathbf{h}_p) \Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p)\, G_0(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_q)\, G_0(\mathbf{x}_q, \omega_k, \mathbf{x}_s), \tag{78}$$

with

$$\gamma_{\,jq}^{\,k} = F(\omega_k) \frac{4\ \omega_k^4}{m_0(\mathbf{x}_j)^3\ m_0(\mathbf{x}_q)^3}. \tag{79}$$

In order to simplify notation, we used G_0 (and g_0) to denote the Green's function in the frequency domain (and time domain) computed with a velocity model $\mathbf{m}_0 = \mathbf{b}_0 + \tilde{\mathbf{r}}_0$. Equation 75 can be rewritten into the matrix form

$$\Delta\tilde{\mathbf{D}} = \begin{pmatrix} \Delta\tilde{D}_{11} \\ \Delta\tilde{D}_{21} \\ \vdots \\ \Delta\tilde{D}_{N_d1} \\ \vdots \\ \Delta\tilde{D}_{1N} \\ \vdots \\ \Delta\tilde{D}_{N_dN} \end{pmatrix} = \begin{pmatrix} T_{11}^{1} & \cdots & T_{11}^{M} \\ T_{21}^{1} & \cdots & T_{21}^{M} \\ \vdots & \vdots & \vdots \\ T_{N_d1}^{1} & \cdots & T_{N_d1}^{M} \\ \vdots & \vdots & \vdots \\ T_{1N}^{1} & \cdots & T_{1N}^{M} \\ \vdots & \vdots & \vdots \\ T_{N_dN}^{1} & \cdots & T_{N_dN}^{M} \end{pmatrix} \begin{pmatrix} \Delta b(\mathbf{x}_1) \\ \Delta b(\mathbf{x}_2) \\ \vdots \\ \Delta b(\mathbf{x}_M) \end{pmatrix}. \tag{80}$$

Adjoint of the tomographic operator

The adjoint of the tomographic operator, as defined in Almomin (2013) is a linear operator that relates a perturbation in the perturbation of the data $\Delta(\Delta\mathbf{D})$ to a perturbation in the background velocity model $\Delta\mathbf{b}$ (such that $\mathbf{b} = \mathbf{b}_0 + \Delta\mathbf{b}$), while keeping the other parameters unchanged. It can be represented by

$$\Delta\mathbf{b} = \mathbf{T}^*(\mathbf{b}_0, \tilde{\mathbf{r}})\ \Delta(\Delta\mathbf{D}). \tag{81}$$

From equation 80, we can easily find the adjoint of the tomographic operator. We have, $\forall q \in \{1; M\}$,

$$\begin{aligned} \Delta b(\mathbf{x}_q) &= \sum_{i=1}^{N_d}\sum_{k=1}^{N} (T_{ik}^{\,q})^*\ \Delta\tilde{D}_{ik} \qquad (82) \\ &= \sum_{i=1}^{N_d}\sum_{k=1}^{N} L_1^*(\mathbf{x}_i, \mathbf{x}_q, \omega_k)\Delta\tilde{D}_{ik} + L_2^*(\mathbf{x}_i, \mathbf{x}_q, \omega_k)\Delta\tilde{D}_{ik}, \end{aligned}$$

and where $\Delta b(\mathbf{x}_q) \in \mathbb{R}$. Moreover,

$$\begin{aligned} &L_1^*(\mathbf{x}_i, \mathbf{x}_q, \omega_k)\Delta\tilde{D}_{ik} = \qquad (83) \\ &\quad \sum_{j=1}^{M}\sum_{p=-h}^{h} (\gamma_{jq}^{\,k})^*\, G_0^*(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s)\Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p)\, G_0^*(\mathbf{x}_q, \omega_k, \mathbf{x}_j + \mathbf{h}_p)\, G_0^*(\mathbf{x}_q, \omega_k, \mathbf{x}_i)\Delta\tilde{D}_{ik}, \end{aligned}$$

and

$$L_2^*(\mathbf{x}_i, \mathbf{x}_q, \omega_k)\Delta\tilde{D}_{ik} = \quad (84)$$
$$\sum_{j=1}^{M}\sum_{p=-h}^{h}(\gamma_{jq}^{\,k})^*\, G_0^*(\mathbf{x}_q, \omega_k, \mathbf{x}_s)\, G_0^*(\mathbf{x}_q, \omega_k, \mathbf{x}_j - \mathbf{h}_p)\Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p)\, G_0^*(\mathbf{x}_j + \mathbf{h}_p, \omega_k, \mathbf{x}_i)\Delta\tilde{D}_{ik}.$$

We have obtained an expression for the adjoint of the tomographic operator (equations 82, 83, and 84), which relates perturbation in the perturbation of the data to perturbation in the background velocity model, while keeping the reflectivity perturbation unchanged.

Time domain expression of the adjoint of the tomographic operator

To get better insight into the adjoint of the tomographic operator, it is convenient to express our previous results in the time domain. We start by rearranging the first term of the right side of equation 82

$$\sum_{i=1}^{N_d}\sum_{k=1}^{N} L_1^*(\mathbf{x}_i, \mathbf{x}_q, \omega_k)\Delta\tilde{D}_{ik} = \sum_{k=1}^{N} P_{\mathrm{scat}_1}^*(\mathbf{x}_q, \mathbf{x}_s, \omega_k) P_{\mathrm{rec}_1}(\mathbf{x}_q, \omega_k) \quad (85)$$

where

- $P_{\mathrm{scat}_1}(\mathbf{x}_q, \mathbf{x}_s, \omega_k) = \sum_{j=1}^{M}\sum_{p=-h}^{h} P_{\mathrm{src}_1}(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_s, \omega_k)\Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p)\, G_0(\mathbf{x}_q, \omega_k, \mathbf{x}_j + \mathbf{h}_p)$,
- $P_{\mathrm{src}_1}(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_s, \omega_k) = \dfrac{2\omega_k^2}{m_0(\mathbf{x}_j)^3} F(\omega_k)\, G_0(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s)$, and
- $P_{\mathrm{rec}_1}(\mathbf{x}_q, \omega_k) = \dfrac{2\omega_k^2}{m_0(\mathbf{x}_q)^3} \sum_{i=1}^{N_d} G_0^*(\mathbf{x}_q, \omega_k, \mathbf{x}_i)\Delta\tilde{D}_{ik}$.

We can express each wavefield in the time domain by taking the inverse DFT of $\mathrm{P}_{\mathrm{scat}_1}$, $\mathrm{P}_{\mathrm{src}_1}$, and $\mathrm{P}_{\mathrm{rec}_1}$. Therefore, we have

$$\begin{aligned} p_{\mathrm{scat}_1}(\mathbf{x}_q, \mathbf{x}_s, n) &= \mathrm{DFT}^{-1}\big[P_{\mathrm{scat}_1}(\mathbf{x}_q, \mathbf{x}_s, \omega_k)\big] \quad (86)\\ &= \sum_{j=1}^{M}\sum_{p=-h}^{h}\big[p_{\mathrm{src}_1}(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_s, n)\Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p)\big] * g(\mathbf{x}_q, n, \mathbf{x}_j + \mathbf{h}_p, 0; \mathbf{b}_0, \tilde{\mathbf{r}}_0), \end{aligned}$$

$$\begin{aligned} p_{\text{src}_1}(\mathbf{x}_j,\mathbf{h}_p,\mathbf{x}_s,n) &= \text{DFT}^{-1}\big[P_{\text{src}_1}(\mathbf{x}_j,\mathbf{h}_p,\mathbf{x}_s,\omega_k)\big] \\ &= \text{DFT}^{-1}\big[\frac{2\omega_{\text{k}}^2}{m_0(\mathbf{x}_j)^3}F(\omega_{\text{k}})\,G_0(\mathbf{x}_j-\mathbf{h}_p,\omega_k,\mathbf{x}_s)\big] \\ &= \frac{-2}{m_0(\mathbf{x}_j)^3}\,\ddot{f}(n)*g(\mathbf{x}_j-\mathbf{h}_p,n,\mathbf{x}_s,0;\mathbf{b}_0,\tilde{\mathbf{r}}_0), \end{aligned} \tag{87}$$

$$\begin{aligned} p_{\text{rec}_1}(\mathbf{x}_q,n) &= \text{DFT}^{-1}\big[P_{\text{rec}_1}(\mathbf{x}_q,\omega_k)\big] \\ &= \text{DFT}^{-1}\big[\sum_{i=1}^{N_d}\frac{2\omega_{\text{k}}^2}{m_0(\mathbf{x}_q)^3}G_0^*(\mathbf{x}_q,\omega_k,\mathbf{x}_i)\Delta\tilde{D}_{ik}\big] \\ &= \frac{-2}{m_0(\mathbf{x}_q)^3}\sum_{i=1}^{N_d}\ddot{g}_-(\mathbf{x}_q,n,\mathbf{x}_i,0;\mathbf{b}_0,\tilde{\mathbf{r}}_0)*\Delta\tilde{d}(\mathbf{x}_i,n,\mathbf{x}_s;\mathbf{b},\tilde{\mathbf{r}}). \end{aligned} \tag{88}$$

P_{rec_1} is a scaled time convolution between the second time derivative of the anti-causal Green's function and the perturbation of the perturbation of the data. It is the perturbation of the perturbation of the data propagated backward in time.

Finally, using the property derived in equation 121 (appendix), we can show that

$$\sum_{i=1}^{N_d}\sum_{k=1}^{N}L_1^*(\mathbf{x}_i,\mathbf{x}_q,\omega_k)\Delta\tilde{D}_{ik}=\big[p_{\text{scat}_1}\otimes p_{\text{rec}_1}\big](\mathbf{x}_q,\mathbf{x}_s,0;\mathbf{b}_0,\tilde{\mathbf{r}}_0). \tag{89}$$

Therefore, the first term of the right side of equation 82 is the zero-lag time cross-correlation between p_{scat_1} and p_{rec_1}. We perform a similar analysis for the second term of the right side of equation 82

$$\sum_{i=1}^{N_d}\sum_{k=1}^{N}L_2^*(\mathbf{x}_i,\mathbf{x}_q,\omega_k)\Delta\tilde{D}_{ik}=\big[p_{\text{src}_2}\otimes p_{\text{scat}_2}\big](\mathbf{x}_q,\mathbf{x}_s,0;\mathbf{b}_0,\tilde{\mathbf{r}}_0). \tag{90}$$

Similarly, the second term of equation 82 is the zero-lag time cross-correlation between p_{src_2} and p_{scat_2}, which are expressed by

$$p_{\text{src}_2}(\mathbf{x}_q,\mathbf{x}_s,n) = \frac{-2}{m_0(\mathbf{x}_q)^3}\,\ddot{f}(n)*g(\mathbf{x}_q,n,\mathbf{x}_s,0;\mathbf{b}_0,\tilde{\mathbf{r}}_0), \tag{91}$$

$$p_{\text{scat}_2}(\mathbf{x}_q, \mathbf{x}_s, n) = \tag{92}$$
$$\sum_{j=1}^{M} \sum_{p=\text{-}h}^{h} \left[p_{\text{rec}_2}(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_i, n) \Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p) \right] * g_-(\mathbf{x}_q, n, \mathbf{x}_j - \mathbf{h}_p, 0; \mathbf{b}_0, \tilde{\mathbf{r}}_0),$$

and

$$p_{\text{rec}_2}(\mathbf{x}_j, \mathbf{h}_p, n) = \frac{-2}{m_0(\mathbf{x}_j)^3} \sum_{i=1}^{N_d} \ddot{g}_-(\mathbf{x}_j + \mathbf{h}_p, n, \mathbf{x}_i, 0; \mathbf{b}_0, \tilde{\mathbf{r}}_0) * \Delta\tilde{d}(\mathbf{x}_i, n, \mathbf{x}_s; \mathbf{b}, \tilde{\mathbf{r}}), \tag{93}$$

where $\mathrm{p}_{\text{rec}_2}$ corresponds to the perturbation of the perturbation of the data convolved with a scaled second time derivative of the anti-causal Green's function. Therefore, we can now explicitly rewrite equation 82 in the time domain. The background perturbation at any subsurface location $\mathbf{x}_q$, generated by a perturbation of the perturbation in the data is given by

$$\Delta b(\mathbf{x}_q) = \left[p_{\text{scat}_1} \otimes p_{\text{rec}_1} \right] (\mathbf{x}_q, \mathbf{x}_s, 0; \mathbf{b}_0, \tilde{\mathbf{r}}_0) + \left[p_{\text{src}_2} \otimes p_{\text{scat}_2} \right] (\mathbf{x}_q, \mathbf{x}_s, 0; \mathbf{b}_0, \tilde{\mathbf{r}}_0). \tag{94}$$

Physical interpretation of the adjoint of the tomographic operator

In order to get a physical understanding of the adjoint of the tomographic operator, we limit the study to the zero-subsurface offset case, and we consider the scenario where we have

- a single source located at a point $\mathbf{x}_s$ at the surface,
- a single recording point $\mathbf{x}_i$ (and its recorded data) at the surface,
- a known background background model $\mathbf{b}_0$,
- an unknown low wavenumber (i.e., smooth) background velocity anomaly $\Delta\mathbf{b}$, and
- a known reflectivity model $\mathbf{r}$ (i.e., an image) obtained by applying (for instance) the RTM operator using a smooth background model. That is, $\mathbf{r} = \mathbf{r}_0 + \Delta\mathbf{r}$, where $\mathbf{r}_0 = \mathbf{0}$.

Throughout this example, we assume that the reflectivity model is given and is considered a fixed part of the operator. We also assume that the background velocity anomaly is small (in magnitude) relative to the background background. We wish to find the anomaly $\Delta\mathbf{b}$ that needs to be added to the background background $\mathbf{b}_0$ to obtain the correct background model $\mathbf{b} = \mathbf{b}_0 + \Delta\mathbf{b}$. The setup of our example is illustrated in Figure 3, where a smooth background anomaly $\Delta\mathbf{b}$ is embedded into a known $\mathbf{m}_0 = \mathbf{b}_0 + \mathbf{r}$.

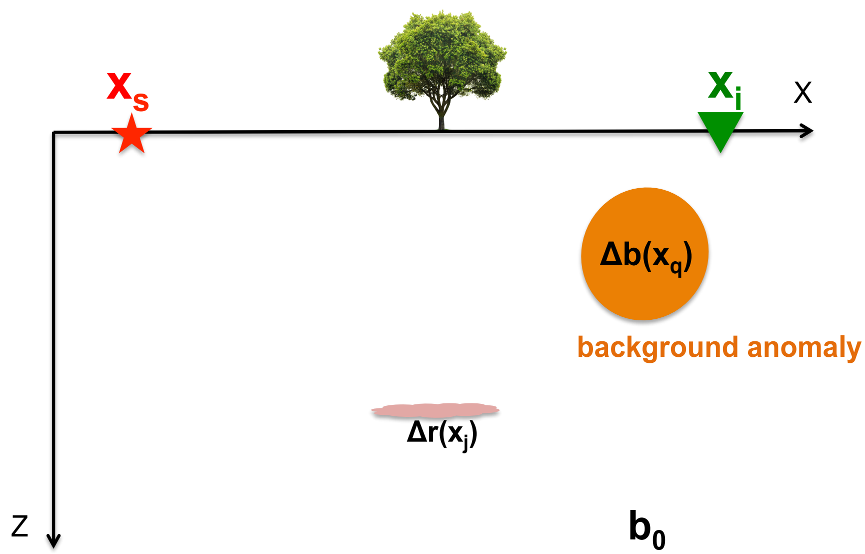


Figure 3: Schematic diagram of the true velocity model used for the experiment. A smooth and relatively small (in magnitude) background anomaly $\Delta\mathbf{b}$ (orange) is embedded into a background velocity model $\mathbf{m}_0$, such that $\mathbf{m}_0 = \mathbf{b}_0 + \mathbf{r}$. The goal is to recover the unknown anomaly $\Delta\mathbf{b}$ by applying the adjoint of the tomographic operator. [**NR**]

In order to recover the background velocity anomaly, we first need to compute the input of the adjoint of the tomographic operator. The perturbation in the perturbation of the data $\Delta\tilde{D}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r})$ recorded at $\mathbf{x}_i$, and due to a seismic source located at $\mathbf{x}_s$, is defined in equation 71 by

$$\begin{aligned}\Delta\tilde{D}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r}) &= \tilde{D}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r}) - \tilde{D}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}_0, \mathbf{r}) \\ &= D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r}) - D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{0}) - \\ &\quad \big(D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}_0, \mathbf{r}) - D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}_0, \mathbf{0})\big).\end{aligned} \quad (95)$$

As mentioned earlier, since we have chosen $\mathbf{r}_0 = \mathbf{0}$, the difference $D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r}) - D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{0})$ corresponds to the field data recorded at the observation point from which we have removed the direct arrivals and diving waves. Similarly, the difference $D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}_0, \mathbf{r}) - D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}_0, \mathbf{0})$ corresponds to the computed data with the background $\mathbf{b}_0$ (which does not contain the anomaly) extracted at the observation point $\mathbf{x}_i$, from which we have removed the direct arrivals and refracted waves. Assuming that we have computed $\Delta\tilde{D}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r})$, we can now interpret the physical meaning of equation 94, starting from the first term of the right side of the equation. The sequence of schematic diagrams in Figure 4 illustrate the following process. The source wavefield $\mathrm{p}_{\mathrm{src}_1}$ is propagated forward in time from location $\mathbf{x}_s$ into the subsurface, with a known velocity model $\mathbf{m}_0 = \mathbf{b}_0 + \mathbf{r}$. In a similar fashion as for the Born modeling operator, a secondary source is created where the source wavefield interacts with a non zero reflectivity perturbation in the subsurface location $\mathbf{x}_j$. This secondary source generates a scattered wavefield $\mathrm{p}_{\mathrm{scat1}}$ (Figure 4(a)). Along with this process, the receiver wavefield $\mathrm{p}_{\mathrm{rec}_1}$ generated by the perturbation in the perturbation of the data $\Delta\tilde{D}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r})$ at location $\mathbf{x}_i$, is propagated backward in time with the known velocity model $\mathbf{m}_0 = \mathbf{b}_0 + \mathbf{r}$ (Figure 4(b)). Finally, $\mathrm{p}_{\mathrm{scat}_1}$ and $\mathrm{p}_{\mathrm{rec}_1}$ are cross-correlated at zero-time lag at every location $\mathbf{x}_q$ in the subsurface to obtain a background velocity perturbation value $\Delta b(\mathbf{x}_q)$ (Figure 4(c)). We can clearly see that the shape of the anomaly coming from the output of the adjoint of the tomographic operator does not correspond to the one of the true anomaly. It is more elongated and less compact.

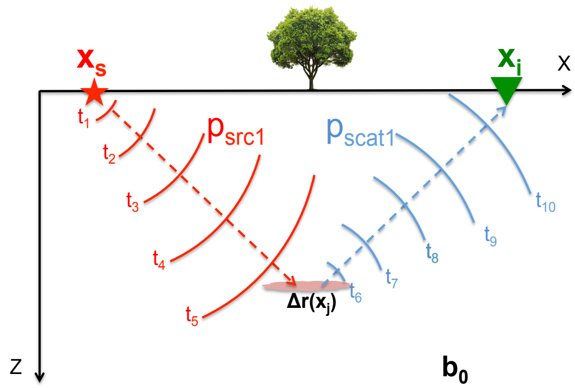
An analogous interpretation can be done for the second term of the right side of equation 94. The source wavefield $\mathrm{p}_{\mathrm{src}_2}$ is propagated forward in time from location $\mathbf{x}_s$ into the subsurface, with a known velocity model $\mathbf{m}_0 = \mathbf{b}_0 + \mathbf{r}$. The receiver wavefield $\mathrm{p}_{\mathrm{rec}_2}$ generated by the perturbation in the perturbation of the data $\Delta\tilde{D}(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r})$ at location $\mathbf{x}_i$, is propagated backward in time with the known velocity model $\mathbf{m}_0 = \mathbf{b}_0 + \mathbf{r}$. When $\mathrm{p}_{\mathrm{rec}_2}$ reaches and interacts with a non zero reflectivity perturbation in the subsurface location $\mathbf{x}_j$, a scattered wavefield $\mathrm{p}_{\mathrm{scat}_2}$ (also propagated backward in time with velocity model $\mathbf{m}_0$) is generated. Finally, $\mathrm{p}_{\mathrm{src}_2}$ and $\mathrm{p}_{\mathrm{scat}_2}$ are cross-correlated at zero-time lag at every location $\mathbf{x}_q$ in the subsurface to obtain a background velocity perturbation value $\Delta b(\mathbf{x}_q)$. The physical process associated with this second term is analogous to the one for the first term. For clarity purposes, it is not illustrated in Figure 4.

Wave-Equation Migration Velocity Analysis (WEMVA) operator

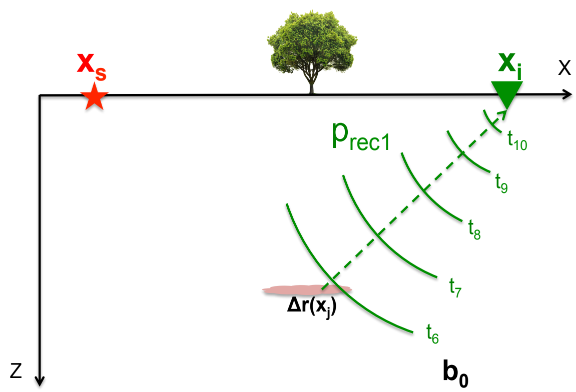
As defined in Almomin (2013), the forward WEMVA operator is an operator that relates a perturbation in the background velocity model $\Delta\mathbf{b}$ (such that $\mathbf{b} = \mathbf{b}_0 + \Delta\mathbf{b}$) to a perturbation in the perturbation of the reflectivity model $\Delta(\Delta\tilde{\mathbf{r}})$, while keeping other parameters unchanged. It is symbolically expressed by

$$\Delta(\Delta\tilde{\mathbf{r}}) = \mathbf{W}(\mathbf{b}_0, \Delta\mathbf{D})\Delta\mathbf{b}. \tag{96}$$

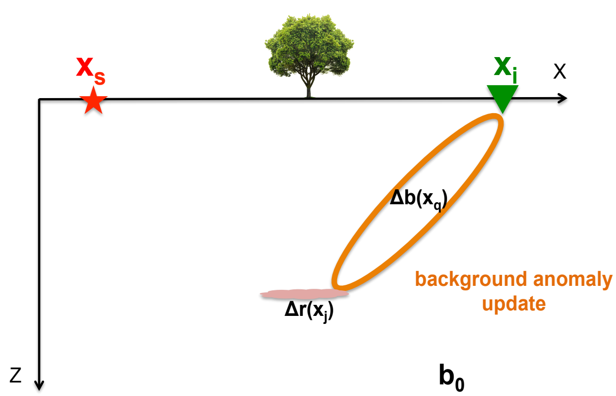
Throughout this process, the data perturbation $\Delta\mathbf{D}$ (expressed in equation 70) is unchanged, and is considered part of the WEMVA operator. It is analogous to the tomographic operator, but instead of performing the linearization on the Born modeling operator around a background background $\mathbf{b}_0$, it is done directly on the RTM operator itself.



(a)



(b)



(c)

Figure 4: Sequence of schematic diagrams illustrating the adjoint of the tomographic operator applied to our example (we only show the effect of the first term in equation 94). (a) The source wavefield interacts with the reflectivity perturbation, acts as a secondary source, and generates the scattered wavefield p_{scat_1}. (b) The receiver wavefield p_{rec_1} is generated by the propagation backward in time of the perturbation of the perturbation in the data. (c) The result of the zero-time lag cross-correlation between p_{scat_1} and p_{rec_1}. [**NR**]

Linearization of the RTM operator

In order to derive the WEMVA operator, we start with the expression of the reflectivity perturbation obtained previously for the RTM operator. The reflectivity perturbation model $\Delta\tilde{\mathbf{r}}$ obtained at all points in the subsurface was computed using a fixed background $\mathbf{b}$. Therefore, the reflectivity perturbation model $\Delta\tilde{\mathbf{r}}$ can also be considered as a function of the background. That is, $\forall j \in \{1; M\}, \forall p \in \{-h; h\}$,

$$\Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p) = \Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p; \mathbf{b}). \tag{97}$$

We can now perform a first-order Taylor expansion of the multivariate function $\Delta\tilde{\mathbf{r}}$ around the background background $\mathbf{b}_0$, which gives

$$\begin{aligned} \Delta(\Delta\tilde{r})(\mathbf{x}_j, \mathbf{h}_p; \mathbf{b}) &= \Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p; \mathbf{b}) - \Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p; \mathbf{b}_0) \\ &\approx \sum_{q=1}^{M} \frac{\partial \Delta\tilde{r}}{\partial b_q}(\mathbf{x}_j, \mathbf{h}_p; \mathbf{b})\Big|_{\mathbf{b}=\mathbf{b}_0} \Delta b(\mathbf{x}_q). \end{aligned} \tag{98}$$

We previously showed (equation 63) that for any given background model $\mathbf{b}$,

$$\begin{aligned} &\Delta\tilde{r}(\mathbf{x}_j, \mathbf{h}_p; \mathbf{b}) = \\ &\quad \sum_{i=1}^{M}\sum_{k=1}^{N} \frac{2F^*(\omega_k)\,\omega_k^2}{m_0(\mathbf{x}_j)^3} G^*(\mathbf{x}_i, \omega_k, \mathbf{x}_j + \mathbf{h}_p; \mathbf{b}, \tilde{\mathbf{r}}_0)\Delta D(\mathbf{x}_i, \omega_k)\, G^*(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s; \mathbf{b}, \tilde{\mathbf{r}}_0). \end{aligned} \tag{99}$$

Now, the term $\Delta D(\mathbf{x}_i, \omega_k)$ is assumed to be known, fixed, and part of the WEMVA operator. Hence,

$$\frac{\partial \Delta\tilde{r}}{\partial b_q}(\mathbf{x}_j, \mathbf{h}_p; \mathbf{x}_s; \mathbf{b}, \tilde{\mathbf{r}})\Big|_{\mathbf{b}=\mathbf{b}_0} = \sum_{i=1}^{M}\sum_{k=1}^{N} \frac{\partial \beta_W}{\partial b_q}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{h}_p, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}_0)\Big|_{\mathbf{b}=\mathbf{b}_0} \alpha_W(\mathbf{x}_j, \omega_k), \tag{100}$$

where

- $\alpha_W(\mathbf{x}_j, \omega_k) = F^*(\omega_k)\dfrac{2\,\omega_k^2}{m_0(\mathbf{x}_j)^3}\Delta D(\mathbf{x}_i, \omega_k)$, and
- $\beta_W(\mathbf{x}_i, \mathbf{x}_j, \mathbf{h}_p, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}_0) = G^*(\mathbf{x}_i, \omega_k, \mathbf{x}_j + \mathbf{h}_p; \mathbf{b}, \tilde{\mathbf{r}}_0)\, G^*(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s; \mathbf{b}, \tilde{\mathbf{r}}_0)$.

The way to evaluate $\frac{\partial \beta_W}{\partial b_q}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{h}_p, \omega_k; \mathbf{b}, \tilde{\mathbf{r}}_0)\Big|_{\mathbf{b}=\mathbf{b}_0}$ is almost identical to the one done for the tomographic operator, and equation 99 can be rewritten in the more compact form

$$\Delta(\Delta\tilde{r})(\mathbf{x}_j, \mathbf{h}_p; \mathbf{b}) = \sum_{q=1}^{M} W_{jp}^{\,q}\,\Delta b(\mathbf{x}_q), \tag{101}$$

$\forall j \in \{1; M\}, \forall p \in \{-h; h\}$. Moreover, $W_{jp}^{\,q} \in \mathbb{C}$ and

$$W_{jp}^{\,q} = L_3(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_q) + L_4(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_q), \tag{102}$$

where

$$\begin{aligned} L_3(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_q) = & \\ & \sum_{i=1}^{N_d}\sum_{k=1}^{N}(\gamma_{\,jq}^{\,k})^* \;\; G_0^*(\mathbf{x}_i, \omega_k, \mathbf{x}_q)\, G_0^*(\mathbf{x}_q, \omega_k, \mathbf{x}_j + \mathbf{h}_p)\Delta D(\mathbf{x}_i, \omega_k)\, G_0^*(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s), \end{aligned} \tag{103}$$

and

$$\begin{aligned} L_4(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_q) = & \\ & \sum_{i=1}^{N_d}\sum_{k=1}^{N}(\gamma_{\,jq}^{\,k})^* \;\; G_0^*(\mathbf{x}_i, \omega_k, \mathbf{x}_j + \mathbf{h}_p)\Delta D(\mathbf{x}_i, \omega_k)\, G_0^*(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_q)\, G_0^*(\mathbf{x}_q, \omega_k, \mathbf{x}_s), \end{aligned} \tag{104}$$

with

$$\gamma_{\,jq}^{\,k} = F(\omega_k)\frac{4\;\omega_k^4}{m_0(\mathbf{x}_j)^3\; m_0(\mathbf{x}_q)^3}. \tag{105}$$

In order to simplify notations, we use G_0 and g_0 to denote the Green's functions (in the frequency and in the time domain, respectively), computed using a velocity model $\mathbf{m}_0 = \mathbf{b}_0 + \tilde{\mathbf{r}}_0$. Equation 96 can be rewritten into the matrix form

$$\Delta(\Delta\tilde{\mathbf{r}}) = \begin{pmatrix} \Delta(\Delta\tilde{r})(\mathbf{x}_1, \mathbf{h}_1; \mathbf{b}) \\ \Delta(\Delta\tilde{r})(\mathbf{x}_1, \mathbf{h}_2; \mathbf{b}) \\ \vdots \\ \Delta(\Delta\tilde{r})(\mathbf{x}_1, \mathbf{h}_{N_h}; \mathbf{b}) \\ \vdots \\ \Delta(\Delta\tilde{r})(\mathbf{x}_M, \mathbf{h}_1; \mathbf{b}) \\ \vdots \\ \Delta(\Delta\tilde{r})(\mathbf{x}_M, \mathbf{h}_{N_h}; \mathbf{b}) \end{pmatrix} = \begin{pmatrix} W_{11}^1 & \cdots & W_{11}^M \\ W_{12}^1 & \cdots & W_{12}^M \\ \vdots & \vdots & \vdots \\ W_{1N_h}^1 & \cdots & W_{1N_h}^M \\ \vdots & \vdots & \vdots \\ W_{M1}^1 & \cdots & W_{M1}^M \\ \vdots & \vdots & \vdots \\ W_{MN_h}^1 & \cdots & W_{MN_h}^M \end{pmatrix} \begin{pmatrix} \Delta b(\mathbf{x}_1) \\ \Delta b(\mathbf{x}_2) \\ \vdots \\ \Delta b(\mathbf{x}_M) \end{pmatrix}, \quad (106)$$

which gives us the expression for the forward WEMVA operator.

Adjoint of the WEMVA operator

The adjoint of the WEMVA operator is a linear operator that relates a perturbation in the perturbation of the reflectivity model $\Delta(\Delta\tilde{\mathbf{r}})$ to a perturbation in the background velocity model $\Delta\mathbf{b}$, while keeping other parameters unchanged (Almomin, 2013). Symbolically, it is expressed by

$$\Delta\mathbf{b} = \mathbf{W}^*(\mathbf{b}_0, \Delta\mathbf{D})\ \Delta(\Delta\tilde{\mathbf{r}}). \quad (107)$$

Equation 101 is easily adjointable, and we have $\forall q \in \{1; M\}$,

$$\begin{aligned} \Delta b(\mathbf{x}_q) &= \sum_{j=1}^{M} \sum_{p=\text{-}h}^{h} (W_{jp}^{\,q})^* \Delta(\Delta\tilde{r})(\mathbf{x}_j, \mathbf{h}_p; \mathbf{b}) \quad (108) \\ &= \sum_{j=1}^{M} \sum_{p=\text{-}h}^{h} \Big(L_3^*(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_q) + L_4^*(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_q) \Big) \Delta(\Delta\tilde{r})(\mathbf{x}_j, \mathbf{h}_p; \mathbf{b}), \end{aligned}$$

where $\Delta b(\mathbf{x}_q) \in \mathbb{R}$, and

$$\begin{aligned} &L_3^*(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_q) = \quad (109) \\ &\quad \sum_{i=1}^{N_d} \sum_{k=1}^{N} \gamma_{\,jq}^{\,k}\ G_0(\mathbf{x}_j - \mathbf{h}_p, \omega_k, \mathbf{x}_s)\, G_0(\mathbf{x}_q, \omega_k, \mathbf{x}_j + \mathbf{h}_p)\, G_0(\mathbf{x}_q, \omega_k, \mathbf{x}_i) \Delta D^*(\mathbf{x}_i, \omega_k), \end{aligned}$$

and

$$L_4^*(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_q) = \sum_{i=1}^{N_d} \sum_{k=1}^{N} \gamma_{jq}^{\,k}\, G_0(\mathbf{x}_q, \omega_k, \mathbf{x}_s)\, G_0(\mathbf{x}_q, \omega_k, \mathbf{x}_j - \mathbf{h}_p)\, G_0(\mathbf{x}_j + \mathbf{h}_p, \omega_k, \mathbf{x}_i) \Delta D^*(\mathbf{x}_i, \omega_k). \tag{110}$$

We have derived the adjoint of the WEMVA operator (equations 108, 109, and 110), which relates perturbation in the perturbation of the reflectivity $\Delta(\Delta\tilde{\mathbf{r}})$ to perturbation in the background velocity model $\Delta\mathbf{b}$, while keeping the data perturbation $\Delta\mathbf{D}$ unchanged.

Time domain expression of the adjoint of the WEMVA operator

By following a similar approach as the one done for the adjoint of the tomographic operator, we obtain the time domain expression for the background model perturbation

$$\Delta b(\mathbf{x}_q) = \left[p_{\text{scat}_3} \otimes p_{\text{rec}_3}\right](\mathbf{x}_q, \mathbf{x}_s, 0; \mathbf{b}_0, \tilde{\mathbf{r}}_0) + \left[p_{\text{src}_4} \otimes p_{\text{scat}_4}\right](\mathbf{x}_q, \mathbf{x}_s, 0; \mathbf{b}_0, \tilde{\mathbf{r}}_0), \tag{111}$$

where

$$p_{\text{scat}_3}(\mathbf{x}_q, \mathbf{x}_s, n) = \sum_{j=1}^{M} \sum_{p=-h}^{h} \left[p_{\text{src}_3}(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_i, n) \Delta(\Delta\tilde{r})(\mathbf{x}_j, \mathbf{h}_p; \mathbf{b})\right] * g(\mathbf{x}_q, n, \mathbf{x}_j + \mathbf{h}_p, 0; \mathbf{b}_0, \tilde{\mathbf{r}}_0), \tag{112}$$

$$p_{\text{src}_3}(\mathbf{x}_j, \mathbf{h}_p, n) = \frac{-2}{m_0(\mathbf{x}_j)^3}\, \ddot{f}(n) * g(\mathbf{x}_j - \mathbf{h}_p, n, \mathbf{x}_s, 0; \mathbf{b}_0, \tilde{\mathbf{r}}_0), \tag{113}$$

$$p_{\text{rec}_3}(\mathbf{x}_q, \mathbf{x}_s, n) = \frac{-2}{m_0(\mathbf{x}_q)^3} \sum_{i=1}^{N_d} \ddot{g}_-(\mathbf{x}_q, n, \mathbf{x}_i, 0; \mathbf{b}_0, \tilde{\mathbf{r}}_0) * \Delta d(\mathbf{x}_i, \mathbf{x}_s, n; \mathbf{b}, \tilde{\mathbf{r}}), \tag{114}$$

and

$$p_{\text{src}_4}(\mathbf{x}_q, \mathbf{x}_s, n) = \frac{-2}{m_0(\mathbf{x}_q)^3}\, \ddot{f}(n) * g(\mathbf{x}_q, n, \mathbf{x}_s, 0; \mathbf{b}_0, \tilde{\mathbf{r}}_0) \tag{115}$$

$$p_{\mathrm{scat}_4}(\mathbf{x}_q, \mathbf{x}_s, n) = \sum_{j=1}^{M} \sum_{p=-h}^{h} \left[p_{\mathrm{rec}_4}(\mathbf{x}_j, \mathbf{h}_p, \mathbf{x}_i, n) \Delta(\Delta\tilde{r})(\mathbf{x}_j, \mathbf{h}_p; \mathbf{b}) \right] * g_-(\mathbf{x}_q, n, \mathbf{x}_j - \mathbf{h}_p, 0; \mathbf{b}_0, \tilde{\mathbf{r}}_0), \tag{116}$$

and

$$p_{\mathrm{rec}_4}(\mathbf{x}_j, \mathbf{h}_p, n) = \frac{-2}{m_0(\mathbf{x}_j)^3} \sum_{i=1}^{N_d} \ddot{g}_-(\mathbf{x}_j + \mathbf{h}_p, n, \mathbf{x}_i, 0; \mathbf{b}_0, \tilde{\mathbf{r}}_0) * \Delta d(\mathbf{x}_i, \mathbf{x}_s, n; \mathbf{b}, \tilde{\mathbf{r}}). \tag{117}$$

We have derived the time domain expression of the adjoint of the WEMVA operator (equation 111). It relates a perturbation of the perturbation of the reflectivity model $\Delta(\Delta\mathbf{r})$ (which can also be interpreted as a perturbation of the image) to a perturbation of the background velocity model $\Delta\mathbf{b}$.

Physical interpretation of the adjoint of the WEMVA operator

In order to get a physical understanding of the mechanism of the adjoint of the WEMVA operator, we consider a similar scenario as for the tomographic operator where we have

- a single source located at a point $\mathbf{x}_s$ at the surface
- a single recording location $\mathbf{x}_i$ (and its recorded data) at the surface
- a known background background model $\mathbf{b}_0$
- an unknown low wavenumber velocity anomaly $\Delta\mathbf{b}$ that we would like to recover

Throughout this example, we define the data perturbation $\Delta D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r})$ as the field data recorded at observation point $\mathbf{x}_i$, from which we have removed the direct arrivals and diving waves. We will assume $\Delta D(\mathbf{x}_i, \mathbf{x}_s, \omega_k; \mathbf{b}, \mathbf{r})$ to be known and unchanged throughout this example. We wish to find the perturbation $\Delta\mathbf{b}$ that needs to be added to the background background $\mathbf{b}_0$ in order to obtain the correct background model $\mathbf{b} = \mathbf{b}_0 + \Delta\mathbf{b}$. Except for the reflectivity model, the setup of our experiment is identical to the previous example (Figure 3).
In order to recover the background velocity anomaly, we first need to compute the input of the adjoint of the WEMVA operator, which is the perturbation in the perturbation of the reflectivity, $\Delta(\Delta\tilde{\mathbf{r}})$. In this example, we assume that we have already

computed a reflectivity perturbation model $\Delta\tilde{\mathbf{r}}$, but using the incorrect background velocity model $\mathbf{b}_0$. Once $\Delta\tilde{\mathbf{r}}$ has been obtained, there are many options to compute $\Delta(\Delta\tilde{\mathbf{r}})$. We will not discuss these methods in this analysis and we assume that we have already computed $\Delta(\Delta\tilde{\mathbf{r}})$. As equations 94 and 111 indicate, the mechanism to obtain the perturbation of the background model is almost identical to the one for the adjoint of the tomographic operator. There are, however, a few variations:

- $\Delta\tilde{\mathbf{r}}$ used to be part of the tomographic operator (and kept unchanged throughout the process), and is now replaced by $\Delta(\Delta\tilde{\mathbf{r}})$,
- $\Delta(\Delta\tilde{\mathbf{r}})$ is now the "input" of the adjoint of the WEMVA operator, and
- $\Delta\tilde{\mathbf{D}}$ is replaced by $\Delta\mathbf{D}$, and is now part of the WEMVA operator, and kept unchanged during the process.

After taking into account those variations, the rest of the process to compute the update in the background velocity model is identical to the one for the adjoint of the tomographic operator, and we refer the reader to the previous section (Figure 4).

SUMMARY

We presented a detailed derivation for the Born, RTM, tomographic, and WEMVA operators by using a Born approximation and multivariate first-order Taylor expansions. We provided an expression of their forward and adjoint (both in the time and in the frequency domain), as well as a physical interpretation of their mechanism.

ACKNOWLEDGMENTS

We would like to thank the Stanford Exploration Project affiliate companies for financial support. Guillaume Barnier would like to thank Ali Almomin, Carlo Fortini, and Gustavo Alves for their unlimited patience, precious help and advice.

APPENDIX

Cross-correlation at zero time lag

We remind the reader a useful equation that relates the cross-correlation at zero-time lag of two time signals to their DFT.

- Let f and g be two discrete real time signals. We assume $\exists N \in \mathbb{N}$ such that $n \notin \{0; N-1\} \Rightarrow f(n) = g(n) = 0$.

- Let h be the time cross-correlation function between f and g. That is, h = f ⊗ g.
- Let F, G, and H be their respective DFT.

Therefore, $\forall n \in \{0; N-1\}$, $\forall k \in \{0; N-1\}$:

$$h(n) = [f \otimes g](n) \Rightarrow H(\omega_k) = F(\omega_k)^* G(\omega_k) = F(\omega_k) G(\omega_k)^*. \tag{118}$$

The definition of the time cross-correlation h between f and g in the time domain is

$$h(n) = \sum_{k=0}^{N-1} f(k+n)\, g(k) = \sum_{k=0}^{N-1} f(k)\, g(k+n). \tag{119}$$

It is also the inverse DFT of the cross-correlation H expressed in the frequency domain,

$$\begin{aligned} h(n) = \mathrm{DFT}^{-1}[H(\omega_k)](n) &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} H(\omega_k)\; e^{i2\pi\, \frac{k\,n}{N}} \qquad (120) \\ &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} F(\omega_k)^*\, G(\omega_k) e^{i2\pi\, \frac{k\,n}{N}} \\ &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} F(\omega_k)\, G(\omega_k)^* e^{i2\pi\, \frac{k\,n}{N}}. \end{aligned}$$

Therefore, taking the cross-correlation at zero time-lag corresponds to evaluating h at $n = 0$, which gives us

$$[f \otimes g](0) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} F(\omega_k)^*\, G(\omega_k) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} F(\omega_k)\, G(\omega_k)^* = \sum_{k=0}^{N-1} f(k)\, g(k). \tag{121}$$

REFERENCES

Almomin, A., 2013, Accurate implementation of two-way wave-equation operators: SEP-Report, **149**, 281–288.