Seismic Imaging [L. Demanet] – *Partial Notes*

Gene Golub SIAM Summer School – 2011

Thibaut Lienart

Acknowledgements: these notes were also based on Nicholas Maxwell's notes.

Caution: these notes remain rough and some parts might be unclear and/or not completely correct, comments are welcome at thibaut.lienart@student.uclouvain.be.

1 Inverse problem and linearization

1.1 Introduction

1. **data** (seismogram): $d_{r,s,t}$ displacement observed by receiver r, source s and time t.
2. **fit** (synthetic seismogram): $u_s(\mathbf{x}_r, t)$ displacement computed by solving a wave equation.
3. **source term** s: $f_s(\mathbf{x}, t) = \delta(\mathbf{x} - \mathbf{x}_s)p_s(t)$ with p_s the source signature (earthquake, noise, ...)
4. **local wave speed**: $c(\mathbf{x})$ (characterization of medium).

The problem is to try to recover the local wave speed $c(\mathbf{x})$ (gives information on the ground structure) from the recorded data. To do so, we try to adjust a synthetic seismogram corresponding to the solution of a wave equation propagating with speed $c(\mathbf{x})$. We define the least-squares misfit between the displacement recorded (data) and the displacement computed (u):

$$J[c] \quad \doteq \quad \frac{1}{2} \|d_{r,s,t} - u_s(\mathbf{x}_r, t)\|_2^2 \tag{1.1}$$

where u_s comes from solving the wave equation corresponding to a source located at $\mathbf{x}_s$ with source function f_s and wave speed $c(\mathbf{x})$. We thus have the following complete minimization problem:

$$\min_{c(\mathbf{x})} J[c] \quad \text{s.t.} \quad \left\{ \begin{array}{rcl} \Box_{c(\mathbf{x})} u_s(\mathbf{x}, t) & = & f_s(\mathbf{x}, t) \\ u_s(\mathbf{x}, 0) & = & 0 \\ \partial_t u_s(\mathbf{x}, 0) & = & 0 \\ \text{proper boundary conditions} & & \end{array} \right. \tag{1.2}$$

Where we have defined the d'Alembertian or wave operator $\Box_c$ by:

$$\Box_c \quad \equiv \quad c^{-2}\partial_t^2 - \Delta \tag{1.3}$$

1.2 Linearization

The previous minimization problem in c is obviously non-linear which makes it hard to solve. Hence, we will try to approximate c using an iterative approach where each step is a linear minimization problem. We will analyze the first such step before describing later the iteration methods.

Assume a known *background wave speed* $c_0(\mathbf{x})$ (*velocity model*) and define the function V such that:

$$c^{-2}(\mathbf{x}) \quad = \quad c_0^{-2}(\mathbf{x}) + V(\mathbf{x}) \tag{1.4}$$

If we assume that V is "small", we can try to make a development in V and only consider first order terms in V.

Let's first define $\mathcal{F}$ the *solution operator* such that:

$$u_s \quad = \quad \mathcal{F}_s[c] \quad \Longleftrightarrow \quad \Box_c u_s \quad = \quad f_s \tag{1.5}$$

We can then develop $\mathcal{F}_s[c]$ around $c_0(\mathbf{x})$:

$$\mathcal{F}_s[c] \quad = \quad \mathcal{F}_s[c_0] + \left(\frac{\delta \mathcal{F}_s}{\delta(c^{-2})}(c_0) \right) \underbrace{[c^{-2} - c_0^{-2}]}_{V} + \text{h.o.t.} \tag{1.6}$$

Where h.o.t. stands for "higher order terms" (in V). Note that the Fréchet derivative of the solution operator taken about c_0 i.e.: $\frac{\delta \mathcal{F}_s}{\delta(c^{-2})}(c_0)$ is itself an operator (*forward operator*). We'll denote this new operator by $\mathbf{F}_s(c_0)$. From now on, we'll call $\mathcal{F}_s[c_0]$ the *incident wave* (for a reason that will appear clear later on) and we'll assume that c_0 is smooth (i.e.: it contains only the low frequency part of c) or in a similar way that V is oscillatory and/or localised (i.e.: contains the high frequency part of c). In that context, we call $\mathbf{F}_s(c_0)[V]$ the *primary reflected wave*, higher order terms would then correspond to a multiply scattered wave. If the condition of smoothness of c_0 is not met then the h.o.t. are significant, otherwise not.[1]

With $u_s = \mathcal{F}_s[c]$ and $u_{\text{inc},s} \doteq \mathcal{F}_s[c_0]$, we can define $u_{\text{scat},s}$ such that:

$$u_s \quad = \quad u_{\text{inc},s} + u_{\text{scat},s} \tag{1.7}$$

with thus $u_{\text{scat},s} = (\mathbf{F}_s(c_0)[V] + \text{h.o.t.})$. We can then define $u^{\text{B}}_{\text{scat},s} = \mathbf{F}_s(c_0)[V]$ (with B for Born, superscript that will be made clear later). Also, note that the definition of $\Box_c$ and the one of V are such that:

$$\Box_c \quad \equiv \quad \Box_{c_0} + V\partial_t^2 \tag{1.8}$$

The wave equation for u_s can hence be rewritten:

$$\Box_{c_0} u_{\text{inc},s} + \Box_{c_0} u_{\text{scat},s} + V\partial_t^2 u_{\text{inc},s} + V\partial_t^2 u_{\text{scat},s} \quad = \quad f_s \tag{1.9}$$

Now, $u_{\text{scat},s}$ is by definition equal to $(\mathbf{F}_s(c_0)[V] + \text{h.o.t.})$ and hence $V\partial_t^2 u_{\text{scat},s}$ is itself a h.o.t.. Comparing the terms of same order, we have:

$$\text{ord.}(1): \qquad \Box_{c_0} u_{\text{inc},s} \quad = \quad f_s \tag{1.10}$$

$$\text{ord.}(V): \qquad \Box_{c_0} u^{\text{B}}_{\text{scat},s} \quad = \quad -V\partial_t^2 u_{\text{inc},s} \tag{1.11}$$

[1] If c_0 is smooth, then the incident wave is only refracted and, through V, we can characterizes well the reflections (due to the boundaries between media of different properties (impedance)). This, however, remains an unclear and unrigorous statement. Finding precise conditions under which this development is valid is still an open question.

The first equation justifies the name *incident wave* as the source term is completely expressed in $u_{\mathrm{inc},s}$.

Only considering first order terms in V, the previous minimization problem in c becomes linear in V with

$$J_L[V] = \frac{1}{2}\| \underbrace{d_{r,s,t} - u_{\mathrm{inc},s}(\mathbf{x}_r,t)}_{\tilde{d}_s(\mathbf{x}_r,t)} - u^{\mathrm{B}}_{\mathrm{scat},s}(\mathbf{x}_r,t)\|^2 \quad (1.12)$$

and the complete minimization problem (*Linear Inverse Problem*, LIP):

$$\min_V J_L[V] \quad \text{s.t.} \quad \left\{ \begin{array}{rcl} \Box_{c_0} u^{\mathrm{B}}_{\mathrm{scat},s}(\mathbf{x},t) & = & -V\partial_t^2 u_{\mathrm{inc},s}(\mathbf{x},t) \\ \text{zero IC's} & & \\ \text{proper BC's} & & \end{array} \right. \quad (1.13)$$

Where we have defined $\tilde{d}_s(\mathbf{x}_r,t) \doteq (d_{r,s,t} - u_{\mathrm{inc},s})$ with then $J_L[V] = \frac{1}{2}\left\|\tilde{d}_s - \mathbf{F}_s(c_0)[V]\right\|_2^2$.

1.3 Lippman-Schwinger equation and Born approximation

Let now G be the Green's function of the wave operator associated to c_0, $u_{\mathrm{inc},s}$ can then be expressed as a convolution:

$$u_{\mathrm{inc},s}(\mathbf{x},t) = \int\int G(\mathbf{x};\mathbf{y},t-\tau) f_s(\mathbf{y},\tau)\mathrm{d}\mathbf{y}\mathrm{d}\tau \quad (1.14)$$

Note that, if c_0 is constant, the Green's function is known. In 3D for example, we have

$$G(\mathbf{x};\mathbf{y},t) \propto \frac{\delta\left(t - \|\mathbf{x}-\mathbf{y}\|/c_0\right)}{\|\mathbf{x}-\mathbf{y}\|} \quad (1.15)$$

in 2D (more often the case in seismology) the Green's function is given by a Hankel function of the first kind. Accordingly, we can define the (linear) operator $\mathcal{G}$ such that

$$u_{\mathrm{inc},s} = \mathcal{G}[f_s] \quad (1.16)$$

If we now come back to the original wave equation and using the definition of V, we had:

$$\begin{aligned} f_s &= \Box_c u_s && (1.17) \\ &= \Box_{c_0} u_s + V\partial_t^2 u_s && (1.18) \end{aligned}$$

but with the equation (1.10) for the incident wave, we had: $f_s = \Box_{c_0} u_{\mathrm{inc},s}$ which yields:

$$\Box_{c_0}(u_s - u_{\mathrm{inc},s}) = -V\partial_t^2 u_s \quad (1.19)$$

we can now use the operator $\mathcal{G}$ associated to $\Box_{c_0}$:

$$u_s - u_{\mathrm{inc},s} = -\mathcal{G}[V\partial_t^2 u_s] \quad (1.20)$$

which is nothing but the *Lippman-Schwinger* equation:

$$u_s = u_{\mathrm{inc},s} - \mathcal{G}[V\partial_t^2 u_s] \quad (1.21)$$

Now, developing the R.H.S. in V and taking only first order terms into account, we have:

$$u_s = \underbrace{u_{\mathrm{inc},s} - \mathcal{G}[V\partial_t^2 u_{\mathrm{inc},s}]}_{\text{Born approximation}} + \text{h.o.t.} \tag{1.22}$$

Now recalling that $u_s = u_{\mathrm{inc},s} + u^{\mathrm{B}}_{\mathrm{scat},s} + \text{h.o.t.}$ it yields:

$$\boxed{u^{\mathrm{B}}_{\mathrm{scat},s} = -\mathcal{G}[V\partial_t^2 u_{\mathrm{inc},s}]} \tag{1.23}$$

This explains the use of the superscript B. Note that this expression is equivalent to the one we already had in (1.11) by comparing terms of order V in (1.9).
Note also that this development is equivalent to considering a Born series expansion up to order V (which has the advantage of making appear a condition on V): denoting $\mathcal{I}$ the identity operator, the Lippman-Schwinger equation can be rewritten:

$$\underbrace{\left(\mathcal{I} + \mathcal{G}(V\partial_t^2)\right)}_{\mathcal{A}}[u_s] = u_{\mathrm{inc},s} \tag{1.24}$$

Under the condition that $\left\|G(V\partial_t^2)\right\| < 1$ (weak scattering[2]), we can inverse $\mathcal{A}$ using Born series (extension of: $(1+x)^{-1} = 1 - x + \text{h.o.t.}$ to operators):

$$u_s = u_{\mathrm{inc},s} - \mathcal{G}[V\partial_t^2 u_{\mathrm{inc},s}] + \text{h.o.t.} \tag{1.25}$$

again developing u_s to the first order in V ($u_s = u_{\mathrm{inc},s} + u^{\mathrm{B}}_{\mathrm{scat},s} + \text{h.o.t.}$) yields (1.23).

[2] In practice this condition is too strong and not practical. Note that the norm is undefined here, we can choose one which will define a space in which the series converge.

2 Iterative method

In the previous section, we have just developed once, starting from a velocity model $c_0(\mathbf{x})$ and seeking to find $c(\mathbf{x})$. We could imagine that the initial guess for c_0 is not very good and then seek to update c_0 while trying to approach c however this can be very expansive as we'll see. On the other hand we could assume that c_0 is a (sufficiently) good guess and only try to find the best V such that the c obtained is already satisfactory. The first approach is called *Landweber Iterations* whereas the second one is called *Migration*.
As the Landweber iterations imply multiple migration steps (step to find the best V in a cheap way at a given stage) we will start with that before focusing on the migration step.

Rmk: Landweber iterations do not work well in practice (convergence is not observed). A modification of these iterations leads to an ameliorated method which yields more satisfactory results and is called *Full Waveform Inversion*.

2.1 Landweber iterations

If we take the previous developments an rewrite it for $c_k(\mathbf{x})$ instead of $c_0(\mathbf{x})$, one gets:

$$\mathcal{F}_s[c] \quad = \quad \mathcal{F}_s[c_k] + \mathbf{F}_s(c_k)[V_k] + \text{h.o.t.} \tag{2.1}$$

where $\mathbf{F}_s(c_k)$ is the operator (linear in V) obtained by taking the Fréchet derivative of $\mathcal{F}_s$ about c_k and $V_k = (c^{-2} - c_k^{-2})$.

2.1.1 Starting from the LIP

In the least-squares sense, the minimum of $J_L[V_k]$ is given by solving the normal equation:

$$\arg\min_{V_k} \left\|\tilde{d}_{(k),s} - \mathbf{F}_s(c_k)[V_k]\right\|_2^2 \quad = \quad \sum_s \underbrace{(\mathbf{F}_s^*(c_k)\mathbf{F}_s(c_k))}_{\text{normal op. } \mathbf{H}_s}{}^{-1}\mathbf{F}_s^*(c_k)[\tilde{d}_{(k),s}] \tag{2.2}$$

with $\tilde{d}_{(k),s} = d_s - \mathcal{F}_s[c_k]$. We can use this method recursively (Newton steps) and have:

$$c_{k+1}^{-2} \quad = \quad c_k^{-2} + \sum_s \left(\mathbf{H}_s(c_k)\right)^{-1} \mathbf{F}_s^*(c_k)[\tilde{d}_{(k),s}] \tag{2.3}$$

This requires inverting the normal operator which is potentially difficult (and expansive). We could however try to use a gradient descent instead. We can develop this method from the gradient of the linearized objective or, equivalently, from the gradient of the full non-linear objective developed about c_k, the latter is shown in the next point (the same iteration method is obtained).

2.1.2 Starting from the original problem

If we want to follow a line-search (gradient descent) we can consider the following scheme:

$$c_{k+1}^{-2} \quad = \quad c_k^{-2} - \alpha \frac{\delta J}{\delta(c^{-2})}[c_k] \tag{2.4}$$

$$\quad = \quad c_k^{-2} - \alpha \frac{\delta J}{\delta V_k}[c_k] \tag{2.5}$$

assuming fixed steps here. Recall that J is defined by

$$J[c] = \frac{1}{2}\|d_s - \mathcal{F}_s[c]\|_2^2 \quad (2.6)$$

$$= \frac{1}{2}\sum_s \langle d_s - \mathcal{F}_s[c], d_s - \mathcal{F}_s[c]\rangle_{r,t} \quad (2.7)$$

taking the inner product of L^2. We can now expand the second term around c_k to make the linear operator $\mathbf{F}_s(c_k)$ appear:

$$J[c] = \frac{1}{2}\sum_s \langle d_s - \mathcal{F}_s[c], d_s - \mathbf{F}_s(c_k)[V_k] + \text{h.o.t.}\rangle \quad (2.8)$$

Using the fact that $\langle a, b\rangle' = \langle a', b\rangle + \langle a, b'\rangle$ and the symmetry of the inner product for real functions, we then have (with slightly messy variational calculus):

$$\frac{\delta J}{\delta V_k}[c] = \sum_s \langle d_s - \mathcal{F}_s[c], -\mathbf{F}_s(c_k) + \text{terms in } V_k\rangle \quad (2.9)$$

For the gradient step, this has to be evaluated at $c = c_k$ or equivalently at $V_k = 0$ which yields:

$$\frac{\delta J}{\delta(c^{-2})}[c_k] = \sum_s \langle d_s - \mathcal{F}_s[c_k], -\mathbf{F}_s(c_k)\rangle \quad (2.10)$$

$$= -\sum_s \mathbf{F}_s^*(c_k)[\underbrace{d_s - \mathcal{F}_s[c_k]}_{\tilde{d}_{(k),s}}] \quad (2.11)$$

$$= -\mathbf{F}^*(c_k)\left[\tilde{d}_{(k)}\right] \quad (2.12)$$

where we have defined the matrix of Fréchet derivatives: $\mathbf{F}^*(c_k)[\cdot] = (\mathbf{F}_1^*(c_k)[\cdot], \ldots, \mathbf{F}_{n_s}^*(c_k)[\cdot])$ and $\tilde{d}_{(k)}$ is the vector of residuals $(d_s - \mathcal{F}_s[c_k])$. Putting everything together, this yields the following line search (FWI):

$$\boxed{c_{k+1}^{-2} = c_k^{-2} + \alpha\mathbf{F}^*(c_k)[\tilde{d}_{(k)}]} \quad (2.13)$$

From what precedes, we see that the adjoint $\mathbf{F}^*$ can provide a good tool towards the estimation of c. Two things are to be noticed, first this method can be very expansive as it requires to solve a wave equation for each step and on top of that the objective is in general not convex.

2.2 Migration

Assume that c_0 is a sufficiently good guess of the low frequency part of c we seek, as before, to minimize $J_L[V]$ (we drop the explicit mention of c_0 in the notation of $\mathbf{F}$ as there is no ambiguity here):

$$\min_V \left\|\tilde{d} - \mathbf{F}[V]\right\|_2^2 \quad (2.14)$$

As was said before, the optimal V is obtained by solving the normal equations:

$$(\mathbf{F}^*\mathbf{F})[V] = \mathbf{F}^*[\tilde{d}] \quad (2.15)$$

In order to avoid computing the inverse of the normal operator, we can try to follow one gradient and hope it is close to the optimal V. The gradient of J_L is obtained in a similar way as in the previous point and we have:

$$\tilde{V} = \alpha\mathbf{F}^*[\tilde{d}] \quad (2.16)$$

with α a calibration factor. And we then have our (once- updated) speed starting from our background velocity model:

$$\tilde{c}^{-2} = c_0^{-2} + \tilde{V} \tag{2.17}$$

Rmk: in practice, the normal operator is close to diagonal which connects the whole theory with that of pseudo-differential operators (ΨDO) or microlocal operators.

3 Adjoint-state method

We will now only consider one migration step and seek an expression for the adjoint operator which, as we have seen, is required for the previous iterative methods. Note first with respect to the interpretation that $\mathbf{F}_s$ maps the image space (space of the speed c) to the data space (space of the displacement u) and that the adjoint does, quite naturally, the opposite.

As we've seen, we're looking for an expression of $(\mathbf{F}^*[\tilde{d}])(x)$ assuming we're interested in the step going from c_0 to $\tilde{c}$. By definition of the adjoint, we have a relation between the inner-product in model space and the inner-product in data space. Considering one source in particular, we have:

$$\underbrace{\left\langle \mathbf{F}_s^*[\tilde{d}_s], V \right\rangle_{\mathbf{x}}}_{\text{model space}} = \underbrace{\left\langle \tilde{d}_s, \mathbf{F}_s[V] \right\rangle_{r,t}}_{\text{data space}} \tag{3.1}$$

$$= \sum_r \int \tilde{d}_s(\mathbf{x}_r, t)(\mathbf{F}_s[V])(\mathbf{x}_r, t)\mathrm{d}t \tag{3.2}$$

Now, in order to replace the sum by an integral, we can define the function $\tilde{d}_{\mathrm{ext},s}$ by:

$$\tilde{d}_{\mathrm{ext},s}(\mathbf{x}, t) = \sum_r \tilde{d}_s(\mathbf{x}_r, t)\delta(\mathbf{x} - \mathbf{x}_r) \tag{3.3}$$

which allows us to write:

$$\left\langle \mathbf{F}_s^*[\tilde{d}_s], V \right\rangle_{\mathbf{x}} = \int\int \tilde{d}_{\mathrm{ext},s}(\mathbf{x}, t)\underbrace{(\mathbf{F}_s[V])(\mathbf{x}, t)}_{u^{\mathrm{B}}_{\mathrm{scat},s}}\mathrm{d}\mathbf{x}\mathrm{d}t \tag{3.4}$$

The appearance of the term $u^{\mathrm{B}}_{\mathrm{scat},s}$ will allow us to make the connection with previous results but for that we would like the wave operator to appear. We can define q_s, the *adjoint wave field* such that:

$$\Box_{c_0} q_s(\mathbf{x}, t) = \tilde{d}_{\mathrm{ext},s}(\mathbf{x}, t) \tag{3.5}$$

and we then have:

$$\left\langle \mathbf{F}_s^*[\tilde{d}_s], V \right\rangle_{\mathbf{x}} = \int\int \Box_{c_0} q(\mathbf{x}, t)u^{\mathrm{B}}_{\mathrm{scat},s}(\mathbf{x}, t)\mathrm{d}\mathbf{x}\mathrm{d}t \tag{3.6}$$

we would now like to make the term $\Box_{c_0} u^{\mathrm{B}}_{\mathrm{scat},s}$ appear in order to use the fact that $\Box_{c_0} u^{\mathrm{B}}_{\mathrm{scat},s} = -V\partial_t^2 u_{\mathrm{inc},s}$ which we have proved before.

3.1 Adjoint-state problem

Using integration by parts we see that we can indeed consider that the wave operator is self-adjoint under the condition that the following four boundary terms are zero (arising from the integration by parts of the previous double integral and expanding the wave operator using its definition):

$$\int \left(\frac{1}{c_0^2} \frac{\partial q_s}{\partial t} u^{\mathrm{B}}_{\mathrm{scat},s} \Big|_0^T \right) \mathrm{d}\mathbf{x} \quad , \qquad \int \left(\frac{1}{c_0^2} q_s \frac{\partial u^{\mathrm{B}}_{\mathrm{scat},s}}{\partial t} u^{\mathrm{B}}_{\mathrm{scat},s} \Big|_0^T \right) \mathrm{d}\mathbf{x} \tag{3.7}$$

$$\int_0^T \int_{\partial S} \left(\frac{\partial q_s}{\partial n} u^{\mathrm{B}}_{\mathrm{scat},s} \right) \mathrm{d}S\mathrm{d}t \quad , \qquad \int_0^T \int_{\partial S} \left(q_s \frac{\partial u^{\mathrm{B}}_{\mathrm{scat},s}}{\partial n} \right) \mathrm{d}S\mathrm{d}t \tag{3.8}$$

The two last terms are obviously zero because of finite speed of propagation (the information carried by $u^{\mathrm{B}}_{\mathrm{scat},s}$ cannot reach the space boundaries in finite time).
For the first and second term, we note that, due to BC's, both integrand are zero at time $t = 0$. A sufficient condition for the self-adjointness of the wave operator is then simply to require that

$$\frac{\partial q_s}{\partial t}(\mathbf{x}, T) \quad = \quad q_s(\mathbf{x}, T) \quad = \quad 0 \tag{3.9}$$

Gathering the two parts, we have defined the *adjoint-state problem*:

$$\left\{ \begin{array}{rcl} \Box_{c_0} q_s(\mathbf{x}, T) & = & \tilde{d}_{\mathrm{ext},s} \\ q_s(\mathbf{x}, T) & = & 0 \\ \partial_n q_s(\mathbf{x}, T) & = & 0 \end{array} \right. \tag{3.10}$$

this propagates *backward in time*: we try to find the adjoint wave field by back-propagating the residual data $\tilde{d}$. Note that, in this point of view, the BC's are the same as for u.
Under the conditions (3.9), we now have:

$$\left\langle \mathbf{F}^*_s[\tilde{d}_s], V \right\rangle_{\mathbf{x}} \quad = \quad \int\int q_s(\mathbf{x}, t) \underbrace{\Box_{c_0} u^{\mathrm{B}}_{\mathrm{scat},s}(\mathbf{x}, t)}_{-V \partial_t^2 u_{\mathrm{inc},s}} \mathrm{d}\mathbf{x}\mathrm{d}t \tag{3.11}$$

$$\int \left(\mathbf{F}^*_s[\tilde{d}_s] \right)(\mathbf{x}) V(\mathbf{x}) \mathrm{d}\mathbf{x} \quad = \quad \int \left[\int -q_s(\mathbf{x}, t) \partial_t^2 u_{\mathrm{inc},s}(\mathbf{x}, t) \mathrm{d}t \right] V(\mathbf{x}) \mathrm{d}\mathbf{x} \tag{3.12}$$

comparing both sides yields:

$$\left(\mathbf{F}^*_s[\tilde{d}_s] \right)(\mathbf{x}) \quad = \quad - \int q_s(\mathbf{x}, t) \partial_t^2 u^{\mathrm{B}}_{\mathrm{inc},s}(\mathbf{x}, t) \mathrm{d}t \tag{3.13}$$

$$= \quad \left\langle q_s, -\partial_t^2 u_{\mathrm{inc},s} \right\rangle_t \tag{3.14}$$

we can then take the Fourier transform in time, Parseval's theorem indicates that:

$$\left\langle q_s, -\partial_t^2 u_{\mathrm{inc},s} \right\rangle_t \quad = \quad \left\langle \hat{q}_s, \omega^2 \hat{u}_{\mathrm{inc},s} \right\rangle_\omega \tag{3.15}$$

where $\hat{}$ denotes the Fourier transform in time. We have used the fact that the Fourier transform of a derivative in time is equal to $(i\omega)$ times the Fourier transform itself (the trick is applied twice here yielding a $(-\omega^2)$ factor). We can further expand this last expression which yields the reversed-time migration.

3.2 Reversed-time migration

We can indeed use the fact that

$$q_s(\mathbf{x},t) \quad = \quad \int \left(\overline{G(\mathbf{x};\mathbf{y},\cdot)} \star \tilde{d}_{\text{ext},s}(\mathbf{y},\cdot)\right)(t)\,\mathrm{d}\mathbf{y} \tag{3.16}$$

where $\star$ denotes the convolution operator and $\overline{G}$ is the Green's function associated to $\square_{c_0}$ but in reversed time. Taking the Fourier transform in time yields:

$$\hat{q}_s(\mathbf{x},\omega) \quad = \quad \int \overline{\hat{G}(\mathbf{x};\mathbf{y},\omega)}\,\hat{\tilde{d}}_{\text{ext},s}(\mathbf{y},\omega)\mathrm{d}\mathbf{y} \tag{3.17}$$

Now, because $\tilde{d}_{\text{ext},s}$ is a sum of dirac impulses, the integral collapses to the following sum:

$$\hat{q}_s(\mathbf{x},\omega) \quad = \quad \sum_r \overline{\hat{G}(\mathbf{x};\mathbf{x}_r,\omega)}\,\hat{\tilde{d}}_s(\mathbf{x}_r,\omega) \tag{3.18}$$

Collecting all the bits and pieces, we finally have:

$$\left(\mathbf{F}_s^*[\tilde{d}_s]\right)(\mathbf{x}) \quad = \quad \sum_r \int \omega^2 \overline{\hat{G}(\mathbf{x};\mathbf{x}_r,\omega)}\,\hat{\tilde{d}}_s\,(\mathbf{x}_r,\omega)\overline{\hat{u}_{\text{inc},s}(\mathbf{x},\omega)}\mathrm{d}\omega \tag{3.19}$$

Now that we have developed $\hat{q}_s$ in terms of the Green's function, we can apply exactly the same procedure for $\hat{u}_{\text{inc},s}$ with:

$$\hat{u}_{\text{inc},s}(\mathbf{x},\omega) \quad = \quad \int \hat{G}(\mathbf{x};\mathbf{y},\omega)\hat{f}(\mathbf{y},\omega)\mathrm{d}\mathbf{y} \tag{3.20}$$

but, by definition, $f_s(\mathbf{x},t) = \delta(\mathbf{x}-\mathbf{x}_s)p_s(t)$ which yields:

$$\hat{u}_{\text{inc},s}(\mathbf{x},\omega) \quad = \quad \hat{G}(\mathbf{x};\mathbf{x}_s,\omega)\hat{p}_s(\omega) \tag{3.21}$$

with eventually:

$$\left(\mathbf{F}^*[\tilde{d}]\right)(\mathbf{x}) \quad = \quad \sum_{r,s} \int \left(\overline{\hat{G}(\mathbf{x};\mathbf{x}_s,\omega)\hat{G}(\mathbf{x};\mathbf{x}_r,\omega)}\right)\omega^2\overline{\hat{p}_s(\omega)}\,\hat{\tilde{d}}_s(\mathbf{x}_r,\omega)\mathrm{d}\omega \tag{3.22}$$

Executing this operation fully is called *reversed-time migration* (RTM).The difficulty in this method is the computation of the Green's functions which can be expansive; a simplified version brings us to the next point: Kirchhoff migration.

3.3 Kirchhoff migration

In order to avoid the full computation of $\hat{G}$, one can try to approach it with the following expression (WKB approximation, valid at high frequency, see requirements for V.):

$$\hat{G}(\mathbf{x};\mathbf{y},\omega) \simeq a(\mathbf{x},\mathbf{y},\omega)\exp(i\omega\tau(\mathbf{x},\mathbf{y})) \tag{3.23}$$

where $\tau(\mathbf{x},\mathbf{y})$ denotes the *travel time* from $\mathbf{x}$ to $\mathbf{y}$. This approximation is usually not accurate with respect to amplitude for all frequencies (WKB expansion, a corresponds to a singular development in ω^{-1}) but can be a good estimation in phase (localization of reflectors).

4 Relation with radar & CT

In the description of the migration step, we have used the fact that:

$$\left\langle \mathbf{F}^*[\tilde{d}], V \right\rangle_{\mathbf{x}} = \left\langle \tilde{d}, \mathbf{F}[V] \right\rangle_{r,s,t} \tag{4.1}$$

Using what precedes (and the fact that V is real), the LHS can be expanded as follow:

$$\left\langle \mathbf{F}^*[\tilde{d}], V \right\rangle_{\mathbf{x}} = \sum_{r,s} \int\int V(x) \left(\overline{\hat{G}(\mathbf{x};\mathbf{x}_s,\omega)\hat{G}(\mathbf{x};\mathbf{x}_r,\omega)}\right) \omega^2 \overline{\hat{p}_s(\omega)} \hat{\tilde{d}}_s(\mathbf{x}_r,\omega)\mathrm{d}\omega\mathrm{d}x \tag{4.2}$$

Using Parseval's theorem, the RHS can on the other hand be expanded like so:

$$\begin{aligned}
\left\langle \tilde{d}, \mathbf{F}[V] \right\rangle_{r,s,t} &= \left\langle \hat{\tilde{d}}, \widehat{\mathbf{F}[V]} \right\rangle_{r,s,\omega} && (4.3)\\
&= \sum_{r,s} \int \hat{\tilde{d}}_s(\mathbf{x}_r,\omega)\overline{\widehat{\mathbf{F}_s[V]}(\mathbf{x}_r,\omega)}\mathrm{d}\omega && (4.4)
\end{aligned}$$

Comparing the two and reordering the first expression, we get:

$$\begin{aligned}
(\widehat{\mathbf{F}_s[V]})(\mathbf{x}_r,\omega) &= \int \left(\overline{\hat{G}(\mathbf{x};\mathbf{x}_s,\omega)\hat{G}(\mathbf{x};\mathbf{x}_r,\omega)}\right) \omega^2 \hat{p}_s(\omega) V(x)\mathrm{d}x && (4.5)\\
&\simeq \int \exp\left(i\omega(\tau(\mathbf{x},\mathbf{x}_s)+\tau(\mathbf{x},\mathbf{x}_r))\right) A(\mathbf{x},\mathbf{x}_s,\omega)V(\mathbf{x})\mathrm{d}\mathbf{x} && (4.6)
\end{aligned}$$

Where $A(\mathbf{x},\mathbf{x}_s,\omega)$ is equal to $a^*(\mathbf{x},\mathbf{x}_s,\omega)\omega^2\hat{p}_s(\omega)$. Now if we consider that sources and reflectors are at the same location,

$$(\widehat{\mathbf{F}_s[V]})(\mathbf{x}_s,\omega) \simeq \int \exp\left(2i\omega\tau(\mathbf{x},\mathbf{x}_s)\right) A(\mathbf{x},\mathbf{x}_s,\omega)V(\mathbf{x})\mathrm{d}\mathbf{x} \tag{4.7}$$

Let's also assume that the velocity model c_0 is constant, we get:

$$(\widehat{\mathbf{F}_s[V]})(\mathbf{x}_s,\omega) \simeq \int \exp\left(2i\omega \left\|\mathbf{x}_s - \mathbf{x}\right\|/c_0\right) A(\mathbf{x},\mathbf{x}_s,\omega)V(\mathbf{x})\mathrm{d}\mathbf{x} \tag{4.8}$$

As before, we use the trick $\left\langle \tilde{d}, \mathbf{F}[V] \right\rangle_{s,r,t} = \left\langle \mathbf{F}^*[\tilde{d}], V \right\rangle_{\mathbf{x}}$ and Parseval's theorem we get:

$$(\mathbf{F}^*[\tilde{d}])(\mathbf{x}) = \sum_s \int \hat{\tilde{d}}_s(\mathbf{x}_s,\omega)\exp\left(-2i\omega\left\|\mathbf{x}_s-\mathbf{x}\right\|/c_0\right)\overline{A(\mathbf{x},\mathbf{x}_s,\omega)}\mathrm{d}\omega \tag{4.9}$$

The interesting fact is that if one seeks to explicitly express the inverse operator such that $\mathbf{F}^{-1}\mathbf{F} = \mathbf{Id}$, the expression is the same as above but for the amplitude:

$$(\mathbf{F}^{-1}[\tilde{d}])(\mathbf{x}) = \sum_s \int \hat{\tilde{d}}_s(\mathbf{x}_s, \omega) \exp\left(-2i\omega \|\mathbf{x}_s - \mathbf{x}\| / c_0\right) B(\mathbf{x}, \mathbf{x}_s, \omega) \mathrm{d}\omega \tag{4.10}$$

Hence the interest of trying to find an accurate approximation of B when using Kirchhoff migration!
Note: we could also think of a continuum of sources along a parametrized curve, the sums above are then replaced by integrals along this curve.

4.1 Radar

In the context of radar, we have (now in the time domain):

$$(\mathbf{F}_s[V])(\mathbf{x}_s, t) \approx \int \delta\left(t - 2\|\mathbf{x}_s - \mathbf{x}\| / c_0\right) V(\mathbf{x}) \mathrm{d}\mathbf{x} \tag{4.11}$$

using Kirchhoff migration we find an estimation of V (Kirchhoff imaging functional $\mathcal{I}_{\mathrm{KM}}$):

$$\begin{aligned}
\mathcal{I}_{\mathrm{KM}}(\mathbf{x}) &= \sum_s \int \delta\left(t - 2\|\mathbf{x}_s - \mathbf{x}\| / c_0\right) (\mathbf{F}_s[V])(\mathbf{x}_s, t) \mathrm{d}t && (4.12)\\
&= \sum_s (\mathbf{F}_s[V])(\mathbf{x}_s, 2\|\mathbf{x}_s - \mathbf{x}\| / c_0) && (4.13)
\end{aligned}$$

with far field approximation, the norm can be approximated by a dot product.

4.2 CT

In the context of computerized tomography, we have:

$$g(\boldsymbol{\theta}, s) = \int \delta(s - \mathbf{x} \cdot \boldsymbol{\theta}) f(\mathbf{x}) \mathrm{d}\mathbf{x} \tag{4.14}$$

using Kirchhoff migration, we have

$$\begin{aligned}
\mathcal{I}_{\mathrm{KM}} &= \int \delta(s - \mathbf{x} \cdot \boldsymbol{\theta}) g(\boldsymbol{\theta}, s) \mathrm{d}\boldsymbol{\theta} \mathrm{d}s && (4.15)\\
&= \int g(\boldsymbol{\theta}, \mathbf{x} \cdot \boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} && (4.16)
\end{aligned}$$

In seismology,

$$\begin{aligned}
d_{\mathrm{lin}}(r, t) &\approx \int \delta(t - (\tau(r, x) + \tau(s, x))) V(x) \mathrm{d}x && (4.17)\\
\mathcal{I}_{\mathrm{KM}}(x) &= \sum_r \int \delta(t - (\tau(r, x) + \tau(s, x))) d_{\mathrm{lin}}(r, t) \mathrm{d}t && (4.18)\\
&= \sum_r d_{\mathrm{lin}}(r, \tau(r, x) + \tau(s, x)) && (4.19)
\end{aligned}$$

with $\tau(r, x) = \|r - x\| / c$.
The point here is that all three contexts have very similar governing equations, this can actually be placed in the Generalized Radon Transform context and hence the inversion algorithms are similar.

5 WKB approximation for the Green's function

As we have seen for Kirchhoff migration, we use an approximation for the Fourier transform of the Green's function. We will now try to explain this a little further.
Consider the Fourier transform (in time) of the general wave equation:

$$\left(\frac{-\omega^2}{c^2(\mathbf{x})} - \Delta_{\mathbf{x}}\right)\hat{G}(\mathbf{x};\mathbf{y},\omega) \quad = \quad \delta(\mathbf{x}-\mathbf{y}) \tag{5.1}$$

or equivalently:

$$\left(\frac{\omega^2}{c^2(\mathbf{x})} + \Delta_{\mathbf{x}}\right)\hat{G}(\mathbf{x};\mathbf{y},\omega) \quad = \quad -\delta(\mathbf{x}-\mathbf{y}) \tag{5.2}$$

writing $k \doteq \omega/c(\mathbf{x})$ (and $\lambda = 2\pi/k$), we find the classical form of a *Helmholtz equation*:

$$\left(\Delta_{\mathbf{x}} + k^2\right)\hat{G}(\mathbf{x};\mathbf{y},\omega) \quad = \quad -\delta(\mathbf{x}-\mathbf{y}) \tag{5.3}$$

We can now try to find a solution with the form of a the *WKB expansion*:

$$\hat{G}(\mathbf{x};\mathbf{y},\omega) \quad = \quad \exp\left(i\omega\tau(\mathbf{x},\mathbf{y})\right)\underbrace{\sum_{j=0} a_j(\mathbf{x},\mathbf{y})\omega^{-j}}_{a(\mathbf{x},\mathbf{y},\omega)} \tag{5.4}$$

with τ, as before, the travel-time (in the previous developments for KM, we had considered only the term in ω^0 of this expansion). For this development, we make the assumption that ω is large (asymptotic expansion in ω^{-1}). This expansion arises also in other fields such as QM. Now we can plug this expansion in the Helmholtz equation and equate like-powers of ω for the first two terms (slightly cumbersome).
Term in ω^2:

$$(\nabla_{\mathbf{x}}\tau \cdot \nabla_{\mathbf{x}}\tau) \quad = \quad \frac{1}{c^2} \qquad \text{(Eikonal Equation)} \tag{5.5}$$

or equivalently: $\|\nabla_{\mathbf{x}}\tau\|_2 = c^{-1}$. Term in ω:

$$2\nabla_{\mathbf{x}}a_0 \cdot \nabla_{\mathbf{x}}\tau + a_0\Delta_{\mathbf{x}}\tau \quad = \quad 0 \qquad \text{(Transport Equation)} \tag{5.6}$$

5.1 Examples

Take $c(\mathbf{x})$ smooth (otherwise τ is not well defined). For example take $c = c_0$ constant, we can have:

$$\tau(\mathbf{x},\mathbf{y}) \quad = \quad \frac{\|\mathbf{x}-\mathbf{y}\|}{c_0} \tag{5.7}$$

$$\nabla_{\mathbf{x}}\tau(\mathbf{x},\mathbf{y}) \quad = \quad \frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|} \tag{5.8}$$

or also,

$$\tau(\mathbf{x},\mathbf{y}) \quad = \quad \frac{x_1}{c_0} \tag{5.9}$$

$$\nabla_{\mathbf{x}}\tau(\mathbf{x},\mathbf{y}) \quad = \quad \begin{pmatrix} 1/c_0 \\ 0 \end{pmatrix} \tag{5.10}$$

Isophase lines of $\exp(i\omega\tau)$ are the isolevel lines of τ and are described as "wavefronts".
Note also that the wavelength is directly proportional to the speed which implies that if the speed decreases so does the wavelength and the wavefronts are more closely packed together.

5.2 Eikonal equation

We had,

$$\|\nabla_{\mathbf{x}}\tau(\mathbf{x},\mathbf{y})\| = \frac{1}{c(\mathbf{x})} \tag{5.11}$$

we define the *rays* as being the characteristic curves: the curves perpendicular to isolevel curves of τ. Let's now consider a ray $\mathbf{X}(t)$, the propagation of this ray can be expressed by:

$$\dot{\mathbf{X}}(t) = \underbrace{c(\mathbf{X}(t))}_{\text{speed}}\underbrace{\frac{\nabla_{\mathbf{x}}\tau(\mathbf{X}(t),\mathbf{y})}{\|\nabla_{\mathbf{x}}\tau(\mathbf{X}(t),\mathbf{y})\|}}_{\text{direction}} = \nabla_{\mathbf{x}}\tau(\mathbf{X}(t),\mathbf{y})c^2(\mathbf{X}(t)) \tag{5.12}$$

if we take the time derivative of $\tau(\mathbf{X}(t),y)$, we get with the chain rule:

$$\frac{d\tau(\mathbf{X}(t),y)}{dt} = \nabla_{\mathbf{x}}\tau(\mathbf{X}(t),\mathbf{y})\cdot\dot{\mathbf{X}}(t) \tag{5.13}$$

$$= \nabla_{\mathbf{x}}\tau(\mathbf{X}(t),\mathbf{y})\cdot\left(c(\mathbf{X}(t))\frac{\nabla_{\mathbf{x}}\tau(\mathbf{X}(t),\mathbf{y})}{\|\nabla_{\mathbf{x}}\tau(\mathbf{X}(t),\mathbf{y})\|}\right) \tag{5.14}$$

$$= c(\mathbf{X}(t))\,\|\nabla_{\mathbf{x}}\tau(\mathbf{X}(t),\mathbf{y})\| = 1 \tag{5.15}$$

which implies that τ is the travel time (so far we hadn't proved it):

$$\tau(\mathbf{X}(t),\mathbf{y}) - \underbrace{\tau(\mathbf{X}(0),\mathbf{y})}_{=0} = t \tag{5.16}$$

with the second term of the LHS being zero because $\mathbf{X}(0)=\mathbf{y}$ and $\tau(\mathbf{y},\mathbf{y})=0$.
We would now like to find an expression of $\dot{\mathbf{X}}(t)$ that does not depend on τ. Let's therefore define $\mathbf{P}(t)$ a function such that:

$$\mathbf{P}(t) \doteq \nabla_{\mathbf{x}}\tau(\mathbf{X}(t),\mathbf{y}) \tag{5.17}$$

with dimension $[TL^{-1}]$ (inverse of a speed) that we'll call the *slowness*. We have:

$$\dot{\mathbf{P}}(t) = \nabla_{\mathbf{x}}\left(\nabla_{\mathbf{x}}\tau\cdot\dot{\mathbf{X}}\right) \tag{5.18}$$

$$= \nabla_{\mathbf{x}}\left(\nabla_{\mathbf{x}}\tau\cdot\nabla_{\mathbf{x}}\tau c^2\right) \tag{5.19}$$

$$= \frac{1}{2}\nabla_{\mathbf{x}}\left(\|\nabla_{\mathbf{x}}\tau\|^2\right)c^2 \tag{5.20}$$

$$= \frac{1}{2}\nabla_{\mathbf{x}}\left(c^{-2}\right)c^2 \qquad \text{(no more } \tau\text{)} \tag{5.21}$$

$$= -\frac{1}{2}\nabla_{\mathbf{x}}(c^2)c^{-2} = -\frac{1}{2}\nabla_{\mathbf{x}}(c^2)\,\|\mathbf{P}(t)\|^2 \tag{5.22}$$

Gathering everything, we have:

$$\left\{\begin{array}{rcl} \dot{\mathbf{X}}(t) & = & \mathbf{P}(t)c^2(\mathbf{X}(t)) \\ \dot{\mathbf{P}}(t) & = & -\frac{1}{2}\|\mathbf{P}(t)\|^2\,\nabla_{\mathbf{x}}c^2(\mathbf{X}(t)) \\ \tau(\mathbf{X}(t),\mathbf{y}) & = & t \end{array}\right. \tag{5.23}$$

for the wave equation in the time domain, we had

$$\begin{cases} \Box_c G(\mathbf{x},\mathbf{y},t) & = & \delta(\mathbf{x}-\mathbf{y}) \\ G(\mathbf{x},\mathbf{y},t) & \approx & a_0(\mathbf{x},\mathbf{y})\delta(\underbrace{t-\tau(x,y)}_{\phi(t,\mathbf{x},\mathbf{y})}) \end{cases} \tag{5.24}$$

in the Fourier domain (Helmholtz):

$$\begin{cases} \left(\Delta_{\mathbf{x}} + \frac{\omega^2}{c^2(\mathbf{x})}\right)\hat{G}(\mathbf{x},\mathbf{y},\omega) & = & -\delta(\mathbf{x}-\mathbf{y}) \\ \hat{G}(\mathbf{x},\mathbf{y},\omega) & \approx & a_0(\mathbf{x},\mathbf{y})\exp\left(i\omega\tau(\mathbf{x},\mathbf{y})\right) \end{cases} \tag{5.25}$$

Hamilton-Jacobi system:

$$\begin{cases} \frac{1}{c^2}\left\|\partial_t\phi\right\|^2 - \left\|\nabla_{\mathbf{x}}\phi\right\|^2 & = & 0 \\ \phi(t,\mathbf{X}(t),\mathbf{y}) & = & \phi(0,\mathbf{X}(0),\mathbf{y}) \end{cases} \tag{5.26}$$

Eikonal equation

$$\begin{cases} \|\nabla_{\mathbf{x}}\tau(\mathbf{x},\mathbf{y})\| & = & \frac{1}{c(\mathbf{x})} \\ \tau(\mathbf{X}(t),\mathbf{y}) & = & \tau(\mathbf{X}(0),\mathbf{y})+t \end{cases} \tag{5.27}$$

Summary: Hamiltonian system (does not depend on τ),

$$\begin{cases} \dot{\mathbf{X}}(t) & = & \nabla_{\mathbf{P}}H(\mathbf{X}(t),\mathbf{P}(t)) \\ \dot{\mathbf{P}}(t) & = & -\nabla_{\mathbf{x}}H(\mathbf{X}(t),\mathbf{P}(t)) \end{cases} \tag{5.28}$$

with $H(\mathbf{x},\mathbf{p}) = \frac{1}{2}c^2(\mathbf{x})\left\|\mathbf{p}\right\|^2$. And *bicharacteristic curves* correspond to

$$H(\mathbf{X}(t),\mathbf{P}(t)) \quad = \quad H(\mathbf{X}(0),\mathbf{P}(0)) \tag{5.29}$$

Fermat principle (basically, a ray takes locally the fastest path):

$$\inf\left\{T \,:\, \exists\mathbf{X}\,:\,\mathbf{X}(0)=\mathbf{x},\,\mathbf{X}(T)=\mathbf{y},\|\mathbf{X}(t)\| = c(\mathbf{X}(t))\,\forall t: 0<t<T\right\} \quad = \quad \tau(\mathbf{x},\mathbf{y}) \tag{5.30}$$

argument is $\mathbf{X}(t)$ that solves the hamiltonian system for some $\mathbf{P}_0$.

$$\tau(\mathbf{x},\mathbf{y}) \quad = \quad \inf\left\{\underbrace{\int_0^S \mathcal{L}(\mathbf{X}(s),\dot{\mathbf{X}}(s))\mathrm{d}s}_{\text{action}} \,:\, \mathbf{X}(0)=\mathbf{x}\,,\,\mathbf{X}(S)=\mathbf{y}\right\} \tag{5.31}$$

with the Lagrangian being defined as $\mathcal{L}(\mathbf{x},\dot{\mathbf{x}}) = \frac{1}{c(\mathbf{x})}\left\|\dot{\mathbf{x}}\right\|$, integrating, we get $\int_\Gamma c^{-1}(\mathbf{x})\mathrm{d}\ell$ and the arg. of min is $\mathbf{X}(t)$ for some p_0. Euler-Lagrange:

$$\frac{\delta\mathcal{L}}{\delta\mathbf{x}} - \frac{d}{ds}\frac{\delta\mathcal{L}}{\delta\dot{\mathbf{x}}} \quad = \quad 0 \quad \Longleftrightarrow \quad \text{ray equations} \tag{5.32}$$

5.3 Transport equation

We had from equating terms in ω resulting in plugging the WKB expansion in the Helmholtz equation the following transport equation:

$$2\nabla_{\mathbf{x}}\tau \cdot \nabla_{\mathbf{x}} a_0 + a_0\, \Delta_{\mathbf{x}}\tau \quad = \quad 0 \tag{5.33}$$

The term $(\nabla_{\mathbf{x}}\tau \cdot \nabla_{\mathbf{x}} a_0)$ can be considered as a rate of change of a_0 in direction of the ray. Taking the time derivative of a_0 along the ray, we get

$$\begin{aligned} \frac{da_0(\mathbf{X}(t),\mathbf{y})}{dt} &= \nabla_{\mathbf{x}} a_0 \cdot \dot{\mathbf{X}} && (5.34)\\ &= c^2\, \nabla_{\mathbf{x}} a_0 \cdot \nabla_{\mathbf{x}}\tau && (5.35)\\ &= -\frac{1}{2} c^2\, a_0 \Delta_{\mathbf{x}}\tau && (5.36) \end{aligned}$$

where, for the last line, we have used the transport equation. We thus have an ODE for a_0 **along each ray** (this is exactly the method of characteristics). Note that $\Delta_{\mathbf{x}}\tau$ can be expressed through $\mathbf{X}, \mathbf{P}, \partial_{\mathbf{x}_0}\mathbf{X}, \partial_{\mathbf{p}_0}\mathbf{P}$.

Now if we multiply the transport equation by a_0, we get:

$$2a_0\nabla_{\mathbf{x}}\tau \cdot \nabla_{\mathbf{x}} a_0 + a_0^2 \Delta_{\mathbf{x}}\tau \quad = \quad 0 \tag{5.37}$$

which can equivalently be rewritten in terms of a divergence:

$$\nabla_{\mathbf{x}} \cdot (a_0 \nabla_{\mathbf{x}}\tau) \quad = \quad 0 \tag{5.38}$$

we can then use the Green's theorem:

$$\begin{aligned} \int\!\!\int_{\Omega} \nabla_{\mathbf{x}} \cdot \left(a_0^2 \nabla_{\mathbf{x}}\tau\right) \mathrm{d}S &= 0 && (5.39)\\ &= \oint_{\partial\Omega} a_0^2 (\nabla_{\mathbf{x}}\tau \cdot \mathbf{n}) \mathrm{d}\ell && (5.40) \end{aligned}$$

If we consider a "ray tube" (in 2D: two sides are given by rays, two sides by wavefronts), the contribution on both "ray sides" disappears (rays are parallel to $\nabla_{\mathbf{x}}\tau$ hence the normal of these sides is orthogonal to $\nabla_{\mathbf{x}}\tau$).
We then have (with w_1 and w_2 for the two remaining sides):

$$\int_{w_1} a_0^2 \, \|\nabla_{\mathbf{x}}\tau\| \, \mathrm{d}\ell \quad = \quad \int_{w_2} a_0^2 \, \|\nabla_{\mathbf{x}}\tau\| \, \mathrm{d}\ell \tag{5.41}$$

which yields (using the Eikonal equation for the norm of the gradient):

$$\frac{a_0^2(\mathbf{x}_0)}{c(\mathbf{x}_0)} \mathrm{d}\ell_0 \quad = \quad \frac{a_0^2(\mathbf{x})}{c(\mathbf{x})} \mathrm{d}\ell \tag{5.42}$$

which yields respectively in 2D and 3D (applying similar developments):

$$a_0(\mathbf{x}) \quad = \quad a_0(\mathbf{x}_0) \sqrt{\frac{c(\mathbf{x})}{c(\mathbf{x}_0)} \frac{\mathrm{d}\ell_0}{\mathrm{d}\ell}} \tag{5.43}$$

$$a_0(\mathbf{x}) \quad = \quad a_0(\mathbf{x}_0) \sqrt{\frac{c(\mathbf{x})}{c(\mathbf{x}_0)} \frac{\mathrm{d}S_0}{\mathrm{d}S}} \tag{5.44}$$

which is nothing but conservation of energy.

Take for example, in 2D, $c = 1$ and opening-cone geometry: $\frac{\mathrm{d}\ell_0}{\mathrm{d}\ell} = \frac{r_0}{r}$ and hence $a_0(\mathbf{x}) \sim \frac{1}{\sqrt{r}}$. In 3D, same geometry, we have $\frac{\mathrm{d}S_0}{\mathrm{d}S} = \frac{r_0^2}{r^2}$ and $a_0(\mathbf{x}) \sim \frac{1}{r}$.

Now the opposite (2D again), take closing-cone geometry: $a_0(\mathbf{x}) = a_0(\mathbf{x}_0)\sqrt{\frac{\mathrm{d}\ell_0}{\mathrm{d}\ell}}$ and $\frac{\mathrm{d}\ell_0}{\mathrm{d}\ell} = \frac{L}{L-r}$. We have that a_0 is blowing up towards infinity as we go towards the focus point (in the limit of finite frequency).

5.4 Extension principles (Symes)

We had

$$
\begin{aligned}
(\mathbf{F}^*[\tilde{d}])(\mathbf{x}) &= \sum_s (\mathbf{F}_s^*[\tilde{d}_s])(\mathbf{x}) \qquad \text{(stack)} && (5.45)\\
\mathbf{F} &= (\mathbf{F}_1, \ldots, \mathbf{F}_{n_s}) && (5.46)\\
\mathbf{F}^* &= \sum_s \mathbf{F}_s^* && (5.47)
\end{aligned}
$$

The idea is to now define $V_s(\mathbf{x}) = (\mathbf{F}_s^*[\tilde{d}_s])(\mathbf{x})$ i.e.: look at images before summing. If c_0 is sufficiently good then $V_{s_1} \approx V_{s_2}$. This is the origin of a Differential Semblance Optimization (DSO) by Symes:

$$
J_{\mathrm{DSO}}[c_0, (V_1, \ldots, V_{n_s})] = \frac{1}{2}\sum_s \|V_{s+1} - V_s\|^2 + \frac{1}{2}\sum_s \left\|\tilde{d}_s - \mathbf{F}_s[V_s]\right\|_2^2 \qquad (5.48)
$$

and then minimize J_{DSO}. (Second term quantifies "data fidelity"). (Rem: $\mathbf{F}_s$ is unknown precisely and depends on c_0, c_0 being the "best guess" of the real c.

5.5 Travel time tomography

Build a travel functional:

$$
J_{\mathrm{TT}}[c] = \frac{1}{2}\|T_{r,s} - \tau(\mathbf{x}_r, \mathbf{x}_s)\|_2^2 \qquad (5.49)
$$

Minimization problem (with Eikonal equation),

$$
\min_c J_{\mathrm{TT}} \quad \text{s.t.} \quad \|\nabla_{\mathbf{x}}\tau(\mathbf{x}, \mathbf{x}_r)\|_2 = \frac{1}{c(\mathbf{x})} \qquad (5.50)
$$

Rem: sound waves send through earth tend to come back up because of increasing pressure (density does not change much) (bulk modulus changing).

Herglotz-Wiechert (1910). Data: $T(\mathbf{x}_r)$. Horizontal slowness $T'(\mathbf{x}) = p(\mathbf{x})$. $c(\mathbf{z})$ increases with depth (see above).

$$
\begin{aligned}
\tau(p, z) &= \int_0^z \frac{1}{c(z')\sqrt{1 - p^2 c^2(z')}}\mathrm{d}z' && (5.51)\\
T(p) &= 2\tau(p, z(p)) && (5.52)\\
&= 2\int_p^{p_0} \frac{1}{\sqrt{q^2 - p^2}} q\,|Z'(q)|\,\mathrm{d}q && (5.53)
\end{aligned}
$$

linked to "Abel" transform or "1/2 integral".

$$\frac{|Z'(q)|}{2q} = \frac{-1}{\pi q} \int_q^\infty \frac{1}{\sqrt{p^2 - q^2}} T'(p) \mathrm{d}p \tag{5.54}$$

Note that if the assumption "c is decreasing" is not correct, the problem is ill-posed.

More about it: *P. Sheares, Intro to seismology*. The bible is *Aki-Richards, Quantitative seismology*, another great reference is *G. Whitham, Linear & Nonlinear waves*