High resolution seismic imaging: Asymptotic approach

By Jean Virieux
Professeur
UGA

American Petroleum Institute, 1986
Time scales

- **Source time**
  from 0.1 sec to 100 sec (rupture vel.)

- **Wave time**
  from secondes to hours

- **Window time**
  from few secondes to days
Length scales

- Fault length
  200 km for $v_r=2$ km/s
- Discontinuity distance
  from few meters to few 100 kms
- Volume sampling
  from few kms to few 1000 kms
Examples

Record of a far earthquake (Müller and Kind, 1976)

Traces from a oil reservoir (Thierry, 1997)

Seismic imaging is a tough problem on the Moon!

Records on the Moon (meteoritic impact) (Latham et al., 1971)

HR seismig imaging
Motivation

- Delayed travel-time **tomography** source:receiver
  - Fitting times => estimation of times
    => 2-pts rays: millions!

- Full Waveform **Inversion/Migration** src/rec:focal point
  - Same challenge with an order increase in magnitude

- Seismogram **modeling**
  - Handling multiple arrivals and filling gaps (shadows)

Rays ⇔ Geometrical Optics ⇔ Infinite Frequency ⇔ Singularities topology

A link with the Catastrophe Theory (René Thom)
Do we need this **complexity**?
FWI: Ray+Born approach

- Ray tracing efficient only in smooth media
- Updating the velocity structure through iterations induces a dramatic increasing complexity in ray tracing: people stays often (always) with the initial ray tracing once done (Background model~Born)
- Asymptotic theory (or high frequency approximation) introduces a complexity in our finite frequency seismic wave propagation: all scales are there while they are not present in the seismic data

In fluid mechanism:
physics below a given small scale is parametrized (upscaled)
References

- Slawinski, M.A., Seismic waves and rays in elastic media, Handbook of geophysical exploration, seismic, exploration, volume 34, Pergamon, 2003
References (selection)

Outline of this partial overview

- Wavefronts and rays
- Ray equation: non-linear ODE
- Paraxial Ray equation: linear ODE
- Hamilton-Jacobi equation: non-linear PDE
- Conclusion
- Perspectives
Translucid Earth

\[ u(x, t) = A(x) S(t - T(x)) \]

\[ u(x, \omega) = A(x) S(\omega) e^{i\omega T(x)} \]

Travel-time \( T(x) \) and Amplitude \( A(x) \)

\( \omega T(x) \) is sometimes called the phase

Wavefront: \( T(x) = T_0 \)

Diffracting medium:
Wavefront coherence lost

Wavefront preserved

24/02/2016

HR seismig imaging
Eikonal equation

Two simple interpretations of wavefront evolution

Orthogonal trajectories are rays in an isotropic medium

Velocity \( c(x) \)

\[ \frac{\Delta L}{\Delta T} = \frac{\Delta T}{\Delta L} = \frac{1}{c(x)} \rightarrow \nabla_x T(x) = \frac{1}{c(x)} \]

\[
(\nabla_x T(x))^2 = \frac{1}{c^2(x)}
\]

Direction? : abs or square

The orientation of the wavefront could not be guessed from the local information on a specific wavefront

Ray construction

Huighens construction

24/02/2016

HR seismig imaging
Eikonal equation for anisotropic media

\[
(\nabla_x T(x))^2 = \frac{1}{c^2(x, \nabla_x T / |\nabla_x T|)}
\]

Phase velocity \( c(x) \)
Ray (group) velocity \( v(x(t)) \)

\( T = \text{cte} \)

Seismic ray: a 1D curved line \( x(t) \)

Tangent \( \frac{dx}{dt} = \dot{x} = v(t) \) defines the ray velocity

Phase velocity depends on the position and on the orientation \( \vec{n} \).
Transport Equation

Tracing neighboring rays defines a ray tube: variation of amplitude depends on energy flux conservation through sections.

Energy flux same at section one and at section two

$$\Delta \mathcal{E}_1 = A_1^2 dS_1 \Delta T_1 = A_2^2 dS_2 \Delta T_2 = \Delta \mathcal{E}_2$$

$$\rightarrow A_1^2 \nabla T_1 \cdot \vec{n} dS_1 = A_2^2 \nabla T_2 \cdot \vec{n} dS_2$$

$$0 = \iint_{Ray}^{Ray} -A_1^2 \nabla T_1 \cdot \vec{n} dS_1 + A_2^2 \nabla T_2 \cdot \vec{n} dS_2$$

$$0 = \iiint_{Ray}^{Ray} d\text{iv}(A^2 \nabla T) dV \Rightarrow d\text{iv}(A^2 \nabla T) = 0$$

$$A(x)(2 \nabla A(x) \cdot \nabla T(x) + A(x) \nabla^2 T(x)) = 0$$

$$2 \nabla A(x) \cdot \nabla T(x) + A(x) \nabla^2 T(x) = 0$$
Transport equation for anisotropic media

Isotropic case: scalar functions

\[ 2 \nabla A(x). \nabla T(x) + A(x) \nabla^2 T(x) = 0 \]

\[ \sum_{i=1}^{3} 2 \frac{\partial A}{\partial x_i} \frac{\partial T}{\partial x_i} + A \frac{\partial^2 T}{\partial x_i^2} = 0 \]

Anisotropic case: vectorial functions

\[ \forall i = 1,3 \]

\[ \sum_{j=1}^{3} \sum_{k=1}^{3} \sum_{l=1}^{3} \frac{\partial}{\partial x_j} \left( c_{ijkl} A_l \frac{\partial T}{\partial x_k} \right) + c_{ijkl} \frac{\partial A_k}{\partial x_l} \frac{\partial T}{\partial x_j} = 0 \]

\[ \bar{u}(x, t) = \bar{A}(x) S(t - T(x)) \]

\[ \bar{u}(x, \omega) = \bar{A}(x) S(\omega) e^{i\omega T(x)} \]

Not easy to understand geometrically how energy flows
Ray equation

Evolution of $\vec{x}$ is given by $\frac{d\vec{x}}{ds}$

Evolution of $\nabla T$ is given by $\frac{d\nabla T}{ds}$

Ray equations

Curvature equation

We define the slowness vector $\vec{p} = \nabla T(x)$ and the position $\vec{q} = \vec{x}(s)$ along the ray
Various non-linear ray equations

Curvilinear stepping

\[
\frac{d\vec{q}(s)}{ds} = c(\vec{q})\vec{p}
\]

\[
\frac{d\vec{p}(s)}{ds} = \nabla_{\vec{q}} \frac{1}{c(\vec{q})}
\]

\[
\frac{dT(s)}{ds} = \frac{1}{c(\vec{q})}
\]

Time stepping

\[
\frac{d\vec{q}(t)}{dt} = c^2(\vec{q})\vec{p}
\]

\[
\frac{d\vec{p}(t)}{dt} = c(\vec{q})\nabla \frac{1}{c(\vec{q})}
\]

Particle stepping

The simplest set

\[
\frac{d\vec{q}(\xi)}{d\xi} = \vec{p}
\]

\[
\frac{d\vec{p}(\xi)}{d\xi} = \frac{1}{c(\vec{q})} \nabla \frac{1}{c(\vec{q})}
\]

\[
\frac{dT(\xi)}{d\xi} = \frac{1}{c^2(\vec{q})}
\]

\[
\begin{align*}
    dT &= \frac{1}{c(\vec{q})} ds = \frac{1}{c(\vec{q})^2} d\xi \\
\end{align*}
\]

under the condition of the eikonal

\[
p^2 = \frac{1}{c^2(\vec{q})}
\]

Which ODE to select for numerical solving? Either t or \(\xi\) sampling.

Many analytical solutions (gradient of velocity; gradient of slowness square)
Polarization

Acoustic case: the unitary vector \( \mathbf{g}_1 = c(q)\mathbf{p} \) supports the P wave vibration.

Elastic case: the independent shear vibration will be along two unitary vectors \( \mathbf{g}_2 \) and \( \mathbf{g}_3 \) such that

\[
\frac{d\mathbf{g}_2}{dt} = \mathbf{g}_2 \cdot \nabla_q c \quad \frac{d\mathbf{g}_3}{dt} = \mathbf{g}_3 \cdot \nabla_q c
\]

Isotropic case: shear vibrations are orthogonal to compression vibrations.

It is enough to follow the evolution of the projection of elastic unitary vectors on one Cartesian coordinate: \( \mathbf{e}_z \cdot \mathbf{g}_2 \) and \( \mathbf{e}_z \cdot \mathbf{g}_3 \) (Psencik, perso. Comm.)

Time stepping

\[
\frac{d\mathbf{g}_2}{dt} = \mathbf{g}_2 \cdot \nabla_q c \quad \frac{d\mathbf{g}_3}{dt} = \mathbf{g}_3 \cdot \nabla_q c
\]

Particle stepping

\[
\frac{d\mathbf{g}_2}{d\xi} = \frac{\mathbf{g}_2 \cdot \nabla_q c}{p^2} \quad \frac{d\mathbf{g}_3}{d\xi} = \frac{\mathbf{g}_3 \cdot \nabla_q c}{p^2}
\]

One additional equation for polarization.

Chapman, 2004, p180
How to solve these equations

ODE versus PDE!
ODE

Lagrangian formulation

Wavefront complexity
Methods of characteristics

Differential geometry (Courant & Hilbert, 1966)

- Non-linear ordinary differential equations
- Lagrangian formulation as we integrate along rays

In opposition to Eulerian formulation where we compute (ray) quantities at fixed positions
Properties of these ODEs

- Intrinsic solutions independent of the coordinate system used to solve it
- If dummy variable for velocity, use it as the variable stepping (often x coordinate)
  \[ \nabla_x \frac{1}{c(q_z)} = 0 \Rightarrow p_x = cte \Rightarrow q_x = q_x^0 + \xi p_x \ (1) \]
- Eikonal equation: a good proxy for testing the accuracy of the ray tracing (not enough used)

In 3D: six or seven equations
In 2D: four or five equations

(1): rectilinear motion of a particle along this axis in mechanics
Hamilton’s ray equations

\[ \frac{d\hat{q}(\xi)}{d\xi} = \hat{p} \]
\[ \frac{d\hat{p}(\xi)}{d\xi} = \frac{1}{c(\hat{q})} \nabla \hat{q} \frac{1}{c(\hat{q})} \]

Hamiltonians:

\[ H(\hat{q}, \hat{p}) = \frac{1}{2} (p^2 - \frac{1}{c^2(q)}) \]

\[ \frac{d\hat{q}(\xi)}{d\xi} = \nabla \hat{q} H \]
\[ \frac{d\hat{p}(\xi)}{d\xi} = -\nabla \hat{p} H \]
\[ \frac{dT}{d\xi} = \hat{p} \cdot \nabla \hat{p} H \]

Mechanics: ray tracing as a particular ballistic problem

Hamilton approach suitable for perturbation
(Henri Poincaré en 1907 « Mécanique céleste », Richard Feymann Prix Nobel 1965)

\[ \hat{q}_0 + \delta \hat{q} \]
\[ \hat{p}_0 + \delta \hat{p} \]

\( \delta q \) and \( \delta p \) "small"

Information around the ray

Meaning of the neighborhood zone
Fresnel zone if finite frequency
Any zone depending on your problem GBS
Ray equations for anisotropic media

**Eikonal equation**

\[ p^2 = \frac{1}{c^2(\vec{q}, \vec{p})} \]

Please remember that the phase velocity depends only on the orientation of slowness vector and not on its length.

**Time stepping**

\[
\frac{dq_i(t)}{dt} = \left( p_i + \frac{1}{c^3(q_i, p_i)} \frac{\partial c(q_i, p_i)}{\partial p_i} \right) \left( \sum_k p_k (p_k + \frac{1}{c^3(q_k, p_k)} \frac{\partial c(q_k, p_k)}{\partial p_k}) \right) \\
\frac{dp_i(t)}{dt} = \left( \frac{1}{c(q_i, p_i)} \nabla q_i \frac{1}{c(q_i, p_i)} \right) \left( \sum_k p_i (p_k + \frac{1}{c^3(q_k, p_k)} \frac{\partial c(q_k, p_k)}{\partial p_k}) \right)
\]

**Particle stepping**

\[
\frac{dq_i(\xi)}{d\xi} = p_i + \frac{1}{c^3(q_i, p_i)} \frac{\partial c(q_i, p_i)}{\partial p_i} \\
\frac{dp_i(\xi)}{d\xi} = \frac{1}{c(q_i, p_i)} \nabla q_i \frac{1}{c(q_i, p_i)} \\
\frac{dT(\xi)}{d\xi} = \sum_i p_i (p_i + \frac{1}{c^3(q_i, p_i)} \frac{\partial c(q_i, p_i)}{\partial p_i})
\]

Particle formulation does not show simpler structure than time formulation for numerical point of view.
Reduced Hamilton formulation

Reduction of ray equations from six to four \hspace{1cm} (Cerveny, 2001, p107)

Solving the eikonal equation for obtaining $p_3$ reducing from one unknown gives

$$p_3 = -\mathcal{H}^R(x_1, x_2, x_3, p_1, p_2)$$

which is a non-linear partial differential equation of first order known as static Hamilton-Jacobi equations.

We may select $x_3$ as the stepping variable, reducing once more from one unknown, removing two unknows and therefore two equations are cancelled.

Ray system $(x_1, x_2, p_1, p_2)$ with the evolution $x_3$ with a variable $\mathcal{H}^R$

The evolution parameter $x_3$ could not be monotonic (turning rays with an extremum in variable $x_3$).

In 3D, five equations
In 2D, three equations including travel-time integration equation
Reduced Hamilton formulation (2)

Turning rays leads us to consider centered ray coordinate system and trigonometric functions! (to be avoided … as slow crunching)

This is not intrinsic to ray equations which can be written in any coordinate system (as well as the related paraxial/dynamic ray equations)

- Computational complexity for number crunching

- No evaluation in the vicinity of the hyper-surface of the eikonal conservation as we live in lower-dimension space

What is best!

If you are afraid of curvilinear non-orthogonal coordinate system intrinsically linked with the ray geometry, you may forget about that. Cartesian coordinate system leads to higher dimensions but with simpler expressions (somehow faster …)
Mechanical point of view

A conservative system
\[ \mathcal{H}(\ddot{q}, \dot{p}) = 0 \]

A non-conservative system
\[ \mathcal{H}(\ddot{q}, \dot{p}) = \mathcal{E}(\ddot{q}, \dot{p}) \Rightarrow \mathcal{H}(\ddot{q}, \dot{p}) - \mathcal{E}(\ddot{q}, \dot{p}) = \mathcal{H}'(\ddot{q}, \dot{p}) = 0 \]

Embedding it into an isolated system!

Potential energy
\[ \mathcal{H}(\ddot{q}, \dot{p}) = \frac{1}{2} p^2 + V(\ddot{q}) \]

Kinetic energy
\[ \mathcal{H}(\ddot{q}, \dot{p}) = \frac{1}{2} p^2 - \frac{1}{2} \frac{1}{c^2} c(\ddot{q}) \]

Ray is a circle
\[ c(z) = c_0 (1 + \gamma z) \]

I should draw a 3D curve with the x coordinate: Help!

Just FUN?

Hyper-surface
Ray lives there (and not outside)

Chapman’s egg
If perturbation of the velocity structure, should we reset ray tracing?

\[
\mathcal{H}(\dot{q}, \dot{p}) = \mathcal{H}_0(\dot{q}, \dot{p}) + \Delta \mathcal{H}(\dot{q}, \dot{p})
\]

Rays on the Egg\(_0\) used for the estimation of rays on the new Egg

Shifts in the phase space of both sources and receivers: application to extended image analysis as promoted by different people: W. Symes, P. Sava among others.

We may feel the extra-dimensionality around a given Chapman’s egg in this full Hamiltonian formulation: this is not the case when we consider the reduced Hamiltonian.
Velocity variation \( v(z) \)

Ray equations are

The horizontal component of the slowness vector is constant: the trajectory is inside a plan which is called the plan of propagation. We may define the frame \((xoz)\) as this plane.

\[
\begin{align*}
\frac{dq_x}{d\tau} &= p_x; \\
\frac{dq_y}{d\tau} &= p_y; \\
\frac{dq_z}{d\tau} &= p_z \\
\frac{dp_x}{d\tau} &= 0; \\
\frac{dp_y}{d\tau} &= 0; \\
\frac{dp_z}{d\tau} &= u(z) \frac{du(z)}{dz}
\end{align*}
\]

Where \( p_x \) is a constante

\[
\begin{align*}
\frac{dq_x}{dq_z} &= \frac{p_x}{p_z} = \pm \sqrt{u^2(z) - p_x^2}
\end{align*}
\]

For a ray towards the depth

\[
q_x(z_1, p_{x1}) = q_x(z_0, p_{x0}) + \int_{z_0}^{z_1} \frac{p_x}{\sqrt{u^2(z) - p_x^2}} dz
\]
Velocity variation $v(z)$

At a given maximum depth $z_p$, the slowness vector is horizontal following the equation

$$q_x(z_1, p_{x1}) = q_{x0} + \int_{z_0}^{z_p} \frac{p_x}{\sqrt{u^2(z) - p_x^2}} \, dz + \int_{z_1}^{z_p} \frac{p_x}{\sqrt{u^2(z) - p_x^2}} \, dz$$

$$T(z_1, p_{x1}) = T_0 + \int_{z_0}^{z_p} \frac{u^2(z)}{\sqrt{u^2(z) - p_x^2}} \, dz + \int_{z_1}^{z_p} \frac{u^2(z)}{\sqrt{u^2(z) - p_x^2}} \, dz$$

If we consider a source at the free surface as well as the receiver, we get

In Cartesian frame

$$X(p) = 2 \int_0^{z_p} \frac{p}{\sqrt{u^2(z) - p^2}} \, dz$$

$$T(p) = 2 \int_0^{z_p} \frac{u^2(z)}{\sqrt{u^2(z) - p^2}} \, dz$$

with $p = u \sin i$

In Spherical frame

$$\Delta = 2 \int_{r_p}^{a} \frac{p}{\sqrt{r^2 u(r)^2 - p^2}} \, \frac{dr}{r}$$

$$T = 2 \int_{r_p}^{a} \frac{r^2 u(r)^2}{\sqrt{r^2 u(r)^2 - p^2}} \, \frac{dr}{r}$$

with $p = r u \sin i$
Velocity structure in the Earth

- Radial Structure
System of ray tracing equations

Curvilinear stepping

\[
\frac{d\tilde{q}(s)}{ds} = c(\tilde{q})\tilde{p}
\]
\[
\frac{d\tilde{p}(s)}{ds} = \nabla\tilde{q} \frac{1}{c(\tilde{q})}
\]
\[
\frac{dT(s)}{ds} = \frac{1}{c(\tilde{q})}
\]

Time stepping

\[
\frac{d\tilde{q}(t)}{dt} = c^2(\tilde{q})\tilde{p}
\]
\[
\frac{d\tilde{p}(t)}{dt} = c(\tilde{q})\nabla \frac{1}{c(\tilde{q})}
\]

Particule stepping

The simplest set

\[
\frac{d\tilde{q}(\xi)}{d\xi} = \tilde{p}
\]
\[
\frac{d\tilde{p}(\xi)}{d\xi} = \frac{1}{c(\tilde{q})} \nabla \frac{1}{c(\tilde{q})}
\]
\[
\frac{dT(\xi)}{d\xi} = \frac{1}{c^2(\tilde{q})}
\]

\[
\frac{1}{c(\tilde{q})} ds = \frac{1}{c(\tilde{q})^2} d\xi
\]

under the condition of the eikonal

\[
p^2 = \frac{1}{c^2(\tilde{q})}
\]

Which ODE to select for numerical solving? Either t or \(\xi\) sampling.

Many analytical solutions (gradient of velocity; gradient of slowness square)
Time integration of ray equations

1D sampling of 2D/3D medium: FAST

Runge-Kutta second-order integration
Predictor-Corrector integration stiffness

Initial conditions EASY

Boundary conditions VERY DIFFICULT

Shooting $\delta p$?
Bending $\delta x$?
Continuing $\delta c$?

Save $p$ conditions if possible!

AND FROM TIME TO TIME IT FAILS!

But we need 2 points ray tracing because we have a source and a receiver to be connected in seismology!

24/02/2016

HR seismic imaging
Ray tracing by rays

Ray tracing

Constant angle step
Ray tracing by rays

Two-point ray tracing

Source (0,0)

 Receivers (2,-2), (4,-2),(6,-2)

(depth is considered as negative)

Please note the irregular angle stepping.
Hamilton’s ray equations

\[
\begin{align*}
\frac{d\tilde{q}(\xi)}{d\xi} &= \tilde{p} \\
\frac{d\tilde{p}(\xi)}{d\xi} &= \frac{1}{c(\tilde{q})}\nabla\tilde{q} \frac{1}{c(\tilde{q})}
\end{align*}
\]

\[
H(\tilde{q}, \tilde{p}) = \frac{1}{2}(p^2 - \frac{1}{c^2(\tilde{q})})
\]

Mechanics: ray tracing as a particular ballistic problem

Hamilton approach suitable for perturbation
(Henri Poincaré en 1907 « Mécanique céleste »,
Richard Feymann Prix Nobel 1965)

\[
\begin{align*}
\dot{q}_0 + \delta \dot{q} \\
\dot{p}_0 + \delta \dot{p}
\end{align*}
\]

\(\delta q\) and \(\delta p\) "small"

Information around the ray

Meaning of the neighborhood zone
Fresnel zone if finite frequency
Any zone depending on your problem GBS
ODE resolution

- Runge-Kutta of second order
- Write a computer program for an analytical law for the velocity: take a gradient with a component along x and a component along z

Home work : redo the same thing with a Runge-Kutta of fourth order (look after its definition)
Consider a gradient of the square of slowness
Runge-Kutta integration

Second-order RK integration

Non-linear ray tracing

Second-order euler integration for paraxial ray tracing is enough!

Linear paraxial ray tracing

Propagator technique

Optical Lens technique

\[
\frac{df}{d\xi} = A(f) \\
\]

\[
f^{1/2} = f^0 + \frac{\Delta \xi}{2} A(f^0) \\
f^1 = f^0 + \Delta \xi A\left(f^{1/2}\right)
\]

\[
\frac{d\delta f}{d\xi} = A(f)\delta f \\
\]

\[
\delta f^{1} = \delta f^0 + \Delta \xi A(f^0)\delta f^0
\]
Analytical solutions

Vertical gradient of the velocity

Gradient of the slowness square
Keeping complexity low

- Ray tracing is a fast 1D integration in 2D/3D

- Ray tracing equations as ODEs may sample the medium quite evenly

- They are lagrangian formulation: we follow a point while tracing rays without regarding the density of rays
Keeping complexity low

Solutions: moving from ODEs to PDEs (Osher et al, 2002) for adequate spatial sampling of the wavefront. Grids control the complexity!
Interpolation challenge
Mitigating the interpolation issue

We cannot guarantee that two rays will stay nearby for small shooting angles changes.

Differential approach will be an answer: the perturbed ray will stay nearby the reference ray …
Paraxial ray theory similar to Gauss optics
Hamilton’s ray equations

\[
\begin{align*}
\frac{d\hat{q}(\xi)}{d\xi} &= \hat{p}, \\
\frac{d\hat{p}(\xi)}{d\xi} &= \frac{1}{c(\hat{q})} \nabla \hat{a} \frac{1}{c(\hat{q})}
\end{align*}
\]

\[H(\hat{q}, \hat{p}) = \frac{1}{2} \left( p^2 - \frac{1}{c^2(\hat{q})} \right)\]

\[
\frac{d\hat{q}(\xi)}{d\xi} = \nabla_{\hat{p}} H
\]

\[
\frac{d\hat{p}(\xi)}{d\xi} = -\nabla_{\hat{q}} H
\]

\[
\frac{d\hat{T}}{d\xi} = \hat{p} \cdot \nabla_{\hat{p}} H
\]

\[
y = \begin{pmatrix} \hat{q} \\ \hat{p} \end{pmatrix}
\]

Information around the ray \( y_0 \)

\[
\begin{pmatrix} \hat{q}_0 + \delta\hat{q} \\ \hat{p}_0 + \delta\hat{p} \end{pmatrix}
\]

\( \delta\hat{q} \) and \( \delta\hat{p} \) "small"

\[
\delta y = \begin{pmatrix} \delta\hat{q} \\ \delta\hat{p} \end{pmatrix}
\]
Paraxial Ray theory

\[ \frac{d(\vec{q}_0 + \delta \vec{q})}{d\xi} = \nabla_{\vec{p}_0 + \delta \vec{p}} H(\vec{q}_0 + \delta \vec{q}, \vec{p}_0 + \delta \vec{p}) \]

\[ \frac{d \delta \vec{q}}{d\xi} = \nabla_{\vec{p}_0 - \vec{q}_0} H(\vec{q}_0, \vec{p}_0) \delta \vec{p} + \nabla_{\vec{p}_0 - \vec{q}_0} H(\vec{q}_0, \vec{p}_0) \delta \vec{q} \]

\[ \frac{d}{d\xi} \begin{bmatrix} \delta \vec{q} \\ \delta \vec{p} \end{bmatrix} = \begin{bmatrix} \nabla_{pq} H^0 & \nabla_{pp} H^0 \\ -\nabla_{pq} H^0 & -\nabla_{qp} H^0 \end{bmatrix} \begin{bmatrix} \delta \vec{q} \\ \delta \vec{p} \end{bmatrix} \]

The matrix \( A \) does not depend on quantities from \( \delta y \) but only on quantities from \( y_0 \): LINEAR PROBLEM (SIMPLE)!

Estimation of ray tube: KMAH index tracking and amplitude evaluation

Estimation of taking-off angles: shooting strategy

Seismogram computation
Paraxial Ray theory

- Solutions are coordinate dependent (differential computation)
- Not restricted to the so-called ray-centered coordinate system (Cerveny, 2001)
- Cartesian formulation is much simpler to handle (Virieux & Farra, 1991)
2D simple linear system

\[ \frac{d}{d\xi} \begin{bmatrix} \delta q_x \\ \delta q_z \\ \delta p_x \\ \delta p_z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0.5 \frac{\partial^2 \frac{1}{c^2(q_0)}}{\partial x^2} & 0.5 \frac{\partial^2 \frac{1}{c^2(q_0)}}{\partial x \partial z} & 0 & 0 \\ 0.5 \frac{\partial^2 \frac{1}{c^2(q_0)}}{\partial z \partial x} & 0.5 \frac{\partial^2 \frac{1}{c^2(q_0)}}{\partial z^2} & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta q_x \\ \delta q_z \\ \delta p_x \\ \delta p_z \end{bmatrix} \]

Linear system

More complex anisotropic structure but still straightforward

Four elementary paraxial trajectories

\( \delta y^t(0) = (1,0,0,0) \)
\( \delta y^t(0) = (0,1,0,0) \)
\( \delta y^t(0) = (0,0,1,0) \)
\( \delta y^t(0) = (0,0,0,1) \)

\[ \delta y^t = (\delta q_x, \delta q_z, \delta p_x, \delta p_z) \]

NOT A paraxial RAY!
2D paraxial conditions

$$\delta H (\xi) = \delta H (0) = 0$$

Paraxial rays require other conservative quantities: the perturbation of the Hamiltonian should be zero (or, in other words, the eikonal perturbation is zero)

If working with the reduced Hamiltonian, this is implicitly set!

$$\frac{\partial H}{\partial p_x} \delta p_x + \frac{\partial H}{\partial p_z} \delta p_z + \frac{\partial H}{\partial q_x} \delta q_x + \frac{\partial H}{\partial q_z} \delta q_z = 0$$

Or in the isotropic case

$$p_x \delta p_x + p_z \delta p_z - \frac{1}{2} \frac{\partial 1/c^2(x,z)}{\partial x} \delta q_x - \frac{1}{2} \frac{\partial 1/c^2(x,z)}{\partial z} \delta q_z = 0$$

Two independent solutions

Similar conditions in 3D
Readily deduced for anisotropy
Point source conditions

\[ \delta q_x(0) = \delta q_z(0) = 0 \Rightarrow p_x(0)\delta p_x(0) + p_z(0)\delta p_z(0) = 0 \]

\[ \delta p_x(0) = \alpha p_z(0) \]

\[ \delta p_z(0) = -\alpha p_x(0) \]

This is enough to verify this condition initially

\[ \alpha \text{ arbitrary constant (linear system)} \]

Paraxial solution

\[ \delta y^a(\xi) = \alpha p_z(0) \delta y_3(\xi) - \alpha p_x(0) \delta y_4(\xi) \]

From paraxial trajectories, one can combine them for paraxial rays as long as the perturbation of the Hamiltonian is zero.

For a point source, the parameter \( \alpha \) could be set to an arbitrary small value: this is a derivative or plan tangent computation (Gauss optics)
Plane source conditions

\[ \delta p_x(0) = \delta p_z(0) = 0 \]

\[ \Rightarrow \quad \frac{1}{2} \frac{\partial 1/c^2(x, z)}{\partial x}(0) \delta q_x(0) + \frac{1}{2} \frac{\partial 1/c^2(x, z)}{\partial z}(0) \delta q_z(0) = 0 \]

\[ \delta q_x(0) = \alpha \frac{\partial 1/c^2(x, z)}{\partial z}(0) \]

\[ \delta q_z(0) = -\alpha \frac{\partial 1/c^2(x, z)}{\partial x}(0) \]

\[ \alpha \text{ arbitrary constant (linear system)} \]

Paraxial solution \[ \delta y^b(\xi) = \alpha \frac{\partial 1/c^2(x, z)}{\partial z}(0) \delta y_1(\xi) - \alpha \frac{\partial 1/c^2(x, z)}{\partial x}(0) \delta y_2(\xi) \]

This is enough to verify this condition initially but gradient of velocity at the source could be quite arbitrary

Cerveny’s condition

Chapman’s condition (only z variation)

We combine the first two paraxial ray trajectories.

Two independent paraxial rays in 2D (\( \delta y^a \) and \( \delta y^b \)): point (seismograms) and plane (beams) paraxial rays
In 2D, the determinant \[
\begin{vmatrix}
    p_x(\xi) & \delta q_x^3 & \delta q_z^3 \\
    p_z(\xi) & \delta q_x^4 & \delta q_z^4 \\
    0 & p_x(0) & p_z(0)
\end{vmatrix}
\]
may change sign:

Increment by one the KMAH index as we have crossed a caustic.

If minor determinants do not change sign, this is a plane caustic (add 1 to KMAH). If they change sign as well, this is a point caustic (add 2 to KMAH).

In 3D, the determinant
\[
\begin{vmatrix}
    p_x(\xi) & \delta q_x^4 & \delta q_y^4 & \delta q_z^4 \\
    p_y(\xi) & \delta q_x^5 & \delta q_y^5 & \delta q_z^5 \\
    p_z(\xi) & \delta q_x^6 & \delta q_y^6 & \delta q_z^6 \\
    0 & p_x(0) & p_y(0) & p_z(0)
\end{vmatrix}
\]
may change sign.

\[u(\vec{q}, t) = A(\vec{q}) e^{i \omega T(x)} e^{-i \frac{1}{2} \text{sgn}(\omega) \text{KMAH}}\]
Euler integration for paraxial ray tracing

First Euler integration for paraxial ray tracing is enough!

Linear paraxial ray tracing
Propagator technique
Optical Lens technique

\[
\frac{d\delta f}{d\xi} = A(f^0)\delta f
\]

\[
\delta f^1 = \delta f^0 + \Delta\xi \ A(f^0)\delta f^0
\]
Paraxial trajectories and rays

Paraxial rays have to be understood as a kind of derivative

Paraxial solutions are coordinate-dependent …

A differential form on the Chapman’s egg
Paraxial trajectories and rays

Fine sampling:
True ray traced at an angle of 31° (blue)
Paraxial ray deduced at an angle of 31° (green) from a ray traced at an angle of 30° (red)

Coarse sampling:
True ray traced at an angle of 35° (blue)
Paraxial ray deduced at an angle of 35° (green) from a ray traced at an angle of 30° (red)
Two points ray tracing: the paraxial shooting method

Consider $\Delta x$ the distance between ray point at the free surface and sensor position.

Solve iteratively $\Delta x = \frac{dq_x}{d\theta} \Delta \theta$

The estimation of the derivative $\frac{d\delta q_x}{d\theta}$ is through the point paraxial computation:

$$\frac{dq_x}{d\theta} = \frac{\partial q_x}{\partial p_x(0)} \frac{dp_x}{d\theta}(0) + \frac{\partial q_x}{\partial p_z(0)} \frac{dp_z}{d\theta}(0)$$

Point paraxial condition (orthogonal perturbation to $\vec{p}$)

Paraxial quantities for derivative

24/02/2016 HR seismig imaging
Efficient 2-pts-ray tracing

An unevenly shooting sampling for reaching receivers at a depth of 2 km with offsets 2, 4, 6, 8 km

Thanks to the paraxial information

Optimization approach
Amplitude estimation

Consider $\Delta L$ the distance between the exit point of a ray at the particule time $\xi$ and the paraxial ray running point.

From point paraxial ray $\delta y^a$

$$\frac{\Delta L}{\Delta \theta} \propto \frac{\delta q_x^a(\xi) p_z(\xi) - \delta q_z^a(\xi) p_x(\xi)}{\sqrt{p_x(\xi)^2 + p_z(\xi)^2}}$$

From point paraxial trajectories $\delta y^a 3$ and $\delta y^a 4$

$$\frac{\Delta L}{\Delta \theta} \propto \frac{[\delta q_x 3(\xi) p_z(0) - \delta q_x 4(\xi) p_x(0)] p_z(\xi) - [\delta q_z 3(\xi) p_z(0) - \delta q_z 4(\xi) p_x(0)] p_x(\xi)}{\sqrt{p_x(\xi)^2 + p_z(\xi)^2}}$$

Thanks to the point paraxial trajectory estimation $\delta q_3$ and $\delta q_4$, we may estimate the geometrical spreading $\Delta L/\Delta \theta$ and, therefore, the ray amplitude $A(\xi) \propto \Delta L/\Delta \theta$

Using plane paraxial solutions $\delta y1$ and $\delta y2$, we can construct a beam as the Gaussian beams

\[
\begin{vmatrix}
  p_x(\xi) & \delta q_x 3(\xi) & \delta q_z 3(\xi) \\
  p_z(\xi) & \delta q_x 4(\xi) & \delta q_z 4(\xi) \\
  0 & p_x(0) & p_z(0)
\end{vmatrix}
\]

\[
\frac{\Delta L}{\Delta \theta} \propto \frac{\sqrt{p_x(\xi)^2 + p_z(\xi)^2}}{\sqrt{p_x(0)^2 + p_z(0)^2}}
\]
Polarization

Chapman, 2004, p180

**Acoustic case:** the unitary vector \( \hat{g}_1 = c(\hat{q})\hat{p} \) supports the P wave vibration

**Elastic case:** the independent shear vibration will be along two unitary vectors \( \hat{g}_2 \) and \( \hat{g}_3 \) such that

\[
\frac{d\hat{g}_2}{dt} = \hat{g}_2 \cdot \nabla \hat{q} c \hat{g}_1
\]

\[
\frac{d\hat{g}_3}{dt} = \hat{g}_3 \cdot \nabla \hat{q} c \hat{g}_1
\]

**Time stepping**

**Particule stepping**

\[
\frac{d\hat{g}_2}{d\xi} = \frac{\hat{g}_2 \cdot \nabla \hat{q} c}{p^2} \hat{g}_1
\]

\[
\frac{d\hat{g}_3}{d\xi t} = \frac{\hat{g}_3 \cdot \nabla \hat{q} c}{p^2} \hat{g}_1
\]

It is enough to follow the evolution of the projection of elastic unitary vectors on one Cartesian coordinate: \( \hat{e}_z \cdot \hat{g}_2 \) and \( \hat{e}_z \cdot \hat{g}_2 \)

(Psencik, perso. Comm.)
Seismic attributes

2 PT ray tracing non-linear problem solved, any attribute could be computed along this line:

- Time (for tomography)
- Amplitude (through paraxial ODE integration fast)
- Polarisation, anisotropy and so on

Moreover, we may bend the ray for a more accurate ray tracing less dependent of the grid step (FMM)

Keep values of \( p \) at source and receiver!

Travel time evolution with the grid step: blue for FMM and black for recomputed time
STEP ONE: ray tracing

(Lambaré et al., 1996)
STEP TWO: times and amplitudes

(Lambaré et al., 1996)

TRAVELTIME \( X=2 \) km

AMPLITUDE \( X=2 \) km
STEP THREE: seismograms
Stop and move to the numerical exercise by your own …
Local ODE

Semi-Lagrangian approach

Sampling control
Ray tracing by wavefronts

Evolution over time:
- folding of the wavefront is allowed
  (still a significant curse of complexity!)

Dynamic sampling:
- undersampling of ray fans
- oversampling of ray fans

Keep an « uniform » sampling of the medium by rays
- by tracking the surrounding density of rays
- by estimating through paraxial approach the ray density

How to compute multiple arrivals?

Lambaré et al (1996)
Vinje et al (1993) widely used in NORSAR software
Ray tracing by wavefronts

An ODE is solved at each point of the wavefront while it is spanned.

Keeping the sampling of the medium more or less uniform.

Rays and wavefronts in an homogeneous medium.

(Lambaré et al., 1996)
Ray tracing by wavefronts

Sampling the wavefront is an heavy task in 2D & 3D.

Still better than oversampling through ray tracing by rays.

Example of wavefront evolution in a smooth version of the Marmousi model.

Smoothness is required!

24/02/2016

HR seismic imaging
Wavefront construction [Vinje, Iversen, Gjøystdal, Lambaré, ...]

- Solve for $\mathbf{x}(t, \alpha)$ and $\mathbf{p}(t, \alpha)$. Discretize in $\alpha$ and trace rays for $\alpha_1, \alpha_2, \alpha_3, \ldots$ where $\alpha_j = j\Delta \alpha$.
- Insert new rays adaptively by interpolation when front resolution deteriorates. E.g.:
  
  If $|\mathbf{x}(t_n, \alpha_{j+1}) - \mathbf{x}(t_n, \alpha_j)| \geq tol$ then insert new ray at $\alpha_{j+1/2}$.
- Interpolate traveltime/phase/amplitude onto regular grid.

(Runbord, 2007)
Geometrical optics resolution

No scale as we are at infinite frequency!

Is it a fair assumption while we have finite frequency wave content?

We must proceed down to a given resolution length under which we do not want to decipher the wavefront: wavefront healing related to so-called viscous solution.

Do we need this complexity of tracking seismic wavefronts!
Few attempts

The medium should be smooth enough for avoiding the wavefront tracking …

Usually done once in the initial model when performing imaging procedure …

A better strategy? Yes, for first arrivals; maybe for multiple arrivals.
PDE

Eulerian approach

Grid control
The eikonal equation is a first-order non-linear partial differential equation of static Hamilton-Jacobi equations: Crandall & Lions (1983, 1984) have shown that « viscous » solutions of such equation can be obtained.

« Computing first-arrival times is equivalent to tracking an interface advancing at a local speed normal to itself » from Sethian & Popovici (1999)

- Level-set methods (Osher & Sethian, 1988)
- Fast marching methods (Sethian, 1996a; Sethian and Popovici, 1999) \( O(\text{NLogN}) \)
- Fast sweeping methods (Zhao, 2005) \( O(N) \)
- Discontinuous Finite-Element methods (Li et al, 2008; Yan & Osher, 2011)

\[
|\vec{p}| = \frac{1}{c(x)} = |\nabla_x T(\vec{x})| \quad \vec{x} \in \Omega \quad \text{Using an upwind, monotone and consistent discretization of } |\nabla_{\vec{x}} T(\vec{x})| \\
T(\vec{x}) = T_0(\vec{x}) \quad \vec{x} \in \Gamma \subset \partial \Omega
\]
Geometrical optics

Helmholtz equation

$$\Delta u + \omega^2 n(x)^2 u = 0.$$ 

Write solution on the form

$$u(x) = A(x, \omega) e^{i \omega \phi(x)}.$$
In a 2D layered medium, let us assume $T$ is known at a level $z = \text{cte}$. Compute $\frac{\partial T}{\partial x}$ along $z = \text{cte}$ by a finite difference approximation. Deduce $\frac{\partial T}{\partial z}$ from eikonal.

Let us assume $T$ is known at a level $z = \text{cte}$. The eikonal equation is:

$$\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial z}\right)^2 = \frac{1}{c(z)^2}$$

From $T$ at a depth of $z$, we have been able to estimate $T$ at a depth of $z + dz$.

**EIKONAL SOLVER**

**ONLY FIRST-ARRIVAL!**
Fast marching method (FMM)

Computing $T(x,z)$ is a stationary boundary problem: discretize it on a grid and find an efficient numerical method to solve it.

The solution is updated by following the causality in a sequential way: updated pointwise in the order the solution is strictly increasing (upwind difference scheme and a heap-sort algorithm)

Sharp interfaces are difficult to describe

$O(N \log N)$

where $N$ is the number of grid points in a direction

From Sethian & Podovici (1999)
The ENO or WENO stencil

How to estimate the discrete $|\nabla T|$?
A first-order Godunov upwind difference scheme

$$
\left[ \left( \frac{T_{i,j} - T_{i,j}^{x_{\min}}}{h} \right)^{+} \right]^{2} + \left[ \left( \frac{T_{i,j} - T_{i,j}^{y_{\min}}}{h} \right)^{+} \right]^{2} = \frac{1}{c^{2}(x,z)}
$$

with $T_{i,j}^{x_{\min}} = \min(T_{i-1,j}, T_{i+1,j})$ and $T_{i,j}^{y_{\min}} = \min(T_{i,j-1}, T_{i,j+1})$

and with $(x)^{+} = x$ if $x > 0$ or $(x)^{+} = 0$ if $x \leq 0$

High-order improved ENO or WENO stencils (Liu et al, 1994; Jiang and Shu, 1996)
Fast sweeping method (FSM)

2D case

\[ \left( \frac{\partial T}{\partial x} \right)^2 + \left( \frac{\partial T}{\partial z} \right)^2 = \frac{1}{c^2(x, z)} \]

Gauss-Seidel iterations with alternating direction sweepings are incorporated into same upwind finite difference stencil: complexity in \( \mathcal{O}(N) \).

Computing \( T(x,z) \) is a stationary boundary problem.

Velocity

Travel time

In 2D at least four sweeps are needed and six sweeps in 3D

Iterations are independent of the grid size

From Zhao (2005)

Singularities at the boundary may induce errors, especially for a point source (Luo & Qian, 2010)

Any mesh could be used
Fast sweeping method (FSM)

2D case

\[
(p_x)^2 + (p_z)^2 = \frac{1}{c^2(x, z)}
\]

Any mesh could be used

\[
T_C = T_A + \frac{\vec{AC}.\vec{p}}{AC} AC \quad T_C > T_A
\]

\[
T_C = T_B + \frac{\vec{BC}.\vec{p}}{BC} BC \quad T_C > T_B
\]

Waves travel along AC or BC direction with an apparent slowness which is the projected slowness value from the true slowness vector.

\[
\begin{pmatrix}
\frac{x_C - x_A}{AC} & \frac{z_C - z_A}{AC} \\
\frac{x_C - x_B}{BC} & \frac{z_C - z_B}{BC}
\end{pmatrix}
\begin{pmatrix}
p_x \\
p_z
\end{pmatrix}
= \begin{pmatrix}
\frac{T_C - T_A}{AC} \\
\frac{T_C - T_B}{BC}
\end{pmatrix}
\]

\[
(p_x, p_z) = \begin{pmatrix}
MT_C + N \\
PT_C + Q
\end{pmatrix}
\]

Quadratic equation: real solution \(T_C\) needed!

Han et al (2015)

24/02/2016

HR seismig imaging
Causality and viscosity

2D case

We provide always a value if A and/or B have a value

# ray tracing

\[(MT_C + N)^2 + (PT_C + Q)^2 = \frac{1}{c^2(x, z)}\]

\[(M^2 + P^2)T_C^2 + 2(MN + PQ)T_C + N^2 + Q^2 - \frac{1}{c^2(x, z)} = 0\]

equation (1)

IF \(T_A\) unset and IF \(T_B\) unset, do nothing

IF \(T_A\) set and IF \(T_B\) unset, compute \(T_C\) as if wave comes from A

IF \(T_A\) unset and IF \(T_B\) set, compute \(T_C\) as if wave comes from B

IF \(T_A\) set and IF \(T_B\) set, do

compute roots of equation (1)

if real solution

check if the «backward ray» intersects the segment [AB]

if yes, update \(T_C\) by this value if it is smaller

if no real solution, compute the viscous solution

as wave is coming from A or from B: select the smallest value

enddo
Fast sweeping method (FSM)

2D case

Eight points stencil:

From the possible eight values (if one is set), take the smallest one.

Sweeping technique:

Four sweeping when applying the stencil

Sweep 1: \( i_1 = 1, n_1, 1; i_2 = 1, n_2, 1 \)
Sweep 2: \( i_1 = n_1, 1, -1; i_2 = 1, n_2, 1 \)
Sweep 3: \( i_1 = n_1, 1, -1; i_2 = n_2, 1, -1 \)
Sweep 4: \( i_1 = 1, n_1, 1; i_2 = n_2, 1, -1 \)

Iteration over sweeps until convergence (no more updating of \( T_C \))

We need to setup a set of values which will be fixed:

at the source, \( T = 0 \), for example
The number of iterations depend on the medium structure: one must be aware that the characteristics of the hyperbolic system should be sampled at least once by the sweeping loop.

In seismics or seismology, we have less dramatic configuration than the one shown in this figure.

Few iterations are necessary to achieve convergence.
2D examples

A smooth model constructed from Marmousi

Wavefronts as deduced from travel-time computation

(Luo and Qian, 2008)

Only first-arrival times

See Taillandier et al (1999) for an application to first-arrival time tomography using the adjoint formulation
3D example

\[
\begin{align*}
H(\phi_{x_1}, \ldots, \phi_{x_d}, x) &= 0, & x \in \Omega \setminus \Gamma, \\
\phi(x) &= g(x), & x \in \Gamma \subseteq \Omega,
\end{align*}
\]

Zhang, Zhao and Qian (2005)
Drawback: only first-arrival times!

Still very useful to back-raytracing once we know times: getting the Jacobian matrix

Once traveltime $T$ is computed over the grid for one source, we may backtrace using the gradient of $T$ from any point of the medium towards the source (should be applied from each receiver)

The surface $\{\text{MIN TIME}\}$ is convex as time increases from the source: one solution!

Back to inversion through rays

Could we do better: multi-arrival times and amplitudes?
Eikonal equation
First arrival property

(a) Correct solution
(b) Eikonal equation

Note:
- $\phi \sim$ traveltime of the wave
- $\phi(x)$ = constant (level sets) represent wave fronts
Eikonal equation
Example
Wavefront construction [Vinje, Iversen, Gjøystdal, Lambaré, ...]

- Solve for $x(t, \alpha)$ and $p(t, \alpha)$. Discretize in $\alpha$ and trace rays for $\alpha_1, \alpha_2, \alpha_3, \ldots$ where $\alpha_j = j \Delta \alpha$.
- Insert new rays adaptively by interpolation when front resolution deteriorates. E.g.: If $|x(t_n, \alpha_{j+1}) - x(t_n, \alpha_j)| \geq \text{tol}$ then insert new ray at $\alpha_{j+1/2}$.
- Interpolate traveltime/phase/amplitude onto regular grid.

(Runbord, 2007)
Time error over the grid (0)

Errors through FMM times

Errors through rays deduced after FMM times

NOT THE SAME COLOR SCALE
(factor 100)

Coarser grid for computation

24/02/2016

HR seismic imaging
Time-dependent Hamilton-Jacobi equation

\[ \mathcal{H}(\tilde{x}, \tilde{p}) = \frac{1}{2}\left(p^2 - \frac{1}{c^2(\tilde{x})}\right) = \mathcal{H}(\tilde{x}, \nabla_{\tilde{x}} T) = 0 \text{ for any } \tilde{x} \]

with \( T(\tilde{x}) = T_{\text{obs}}(\tilde{x}) \) for position \( \tilde{x} \) on boundary

Two techniques for making a link between static HJ equations and time-dependent HJ equations

- The level-set idea (Osher, 1993) by adding a dimension to the problem
  Introducing the new function \( \psi \) such that \( \psi_t + \mathcal{H}(\tilde{x}, \nabla_{\tilde{x}} \psi) = 0 \), one can find the solution \( T(\tilde{x}) \) as the zero-level \( \psi_t(\tilde{x}) = 0 \).

- The so-called paraxial formulation (Leung et al, 2004) similar to what we have done for the reduced Hamiltonian. One prefered spatial direction is assumed as the evolution (« time ») direction:

Often, a semi-lagrangian approach is used with such increase in the space dimensions for computer efficiency (Fomel & Sethian, 2002).
Static H-J equation versus dynamic one

\[ |\nabla T(x, y, z)| = \frac{1}{c(x,y,z)} \] Eikonal or static Hamilton-Jacobi equation

\[ \frac{\partial \phi}{\partial t}(x, y, z, t) + c(x, y, z)|\nabla \phi(x, y, z, t)| = 0 \] Dynamic H-J equation

The solution should converge to the static one.

\[ \phi(x, y, z, t) = T(x, y, z) - t \]

We may use numerical approaches which consider PDE of H-J equations
Time-dependent Hamilton-Jacobi equation (2)

We are back to PDE strategy on fixed grids!

This is not an artificial tool but a very deep insight in physics (Goldstein, 1980, chap10).

- Hamilton (1834) has shown the equivalence between Eikonal equations and H-J equations.
- De Broglie & Schrödinger (1926) has made the connection with wave equation.

The action \( S(\vec{x}, \vec{p}, \xi) = \mathcal{H}(\vec{x}, \vec{p}) - \mathcal{E}\xi \) where the evolution variable \( \xi \) acts as a time for the related particle and the energy \( \mathcal{E} \) is zero for rays. The action \( S \) is proportional (at the Planck constant limit) to the phase for a general dispersive wave

\[
\vec{u} = \vec{A} e^{iS(\xi, \vec{x}, \vec{p})}
\]

and one has the « particle-time » dependent H-J equation,

\[
\frac{\partial S(\xi, \vec{x}, \vec{p})}{\partial \xi} + \mathcal{H} \left( \vec{x}, \frac{\partial S}{\partial \xi} \right) = 0
\]
Phase space


Introduce phase space $(\mathbf{x}, \mathbf{p})$, where $\mathbf{p} \in S^{d-1}$ is local ray direction.

Observation: Wavefront is a smooth curve in phase space.

- 2D problems: 1D curve in 3D phase space $(x, y, \theta)$.
- 3D problems: 2D surface in 5D phase space $(x, y, z, \theta, \alpha)$.

Wavefront in phase space sweeps out a smooth surface – the Lagrangian submanifold.
The paraxial strategy

In a 2D medium, we can consider the eikonal \( |\nabla_{\hat{x}} T(\hat{x})| = \frac{1}{c(\hat{x})} \) for sub-vertical rays and, therefore, we may choose the variable \( z \) as the evolution. We have

\[
\frac{\partial T(z, x, p_x)}{\partial z} + H(z, x, p_x) = 0
\]

A 2D level-set motion in the space \((x, p_x)\) as we treat the variable \( z \) as an artificial time variable and we consider the function \( \varphi(z, x, p_x) \). We have

\[
\frac{D \varphi}{Dz} = \frac{\partial \varphi}{\partial z} + \frac{\partial \varphi}{\partial x} \frac{dx}{dz} + \frac{\partial \varphi}{\partial p_x} \frac{dp_x}{dz}
\]

with initial values \( \varphi(0, x, p_x) = x_s \) as the position of the source. The zero level-set intersection will provide \( \varphi(z, x(z), p_x(z)) = 0 \), i.e. rays intersections at constant depth. Travel-times can be deduced by integration as well.

(Leung & Qian & Osher, 2004)
**Application: Marmousi case**

Medium from 4.8 km to 7.2 km in x and from 0 km to 3 km in z

Grid \((x,z) = 384 \times 122\) with a stepping 25 m \(\times\) 25 m

The source is at \(x=6.0\) km and \(z=2.8\) km

(Qian & Leung, 2004; Qian & Leung, 2006)

Superposition of ODE & PDE solutions

\((100 \times 200 \times 122)\) for \((p_x, x, z)\)

Contours for various \(z\)

Finer grid
In a 2D medium, we could consider time as the evolution variable and solve the level-set problem for the function \( \varphi(t, x, z, p_x, p_z) \) defined by

\[
\frac{\partial \varphi}{\partial t} + \frac{dx}{dt} \frac{\partial \varphi}{\partial x} + \frac{dy}{dt} \frac{\partial \varphi}{\partial z} + \frac{dp_x}{dt} \frac{\partial \varphi}{\partial p_x} + \frac{dp_z}{dt} \frac{\partial \varphi}{\partial p_z}
\]

with two initial conditions \( \varphi^1(0, x, z, p_x, p_z) = x \) and \( \varphi^2(0, x, z, p_x, p_z) = z \).

The intersection of the zero level-set values of these two functions gives the phase space solution \( (x, z, p_x, p_z)(t) \) for source (0,0). In fact, we can have all solutions by selection another level-set values other than zero.

It is an eulerian approach in a 5-D space in 2D medium which could be 7-D space for 3D medium. The approach is not restricted to paraxial formulation.

Strategies have been defined to reduce the cost: reduced phase space, semi-lagrangian integration (Fomel & Sethian, 2002; Osher et al, 2002).

*Work in progress*
Purpose of eulerian approach

- Very robust technique for multiple travel-times estimation: grid controls the complexity (challenging problem)

- Estimating amplitudes has also been setup through eulerian approach

- Identification of branches at the grid spacing accuracy (practical in 2D!)

Strategies to be tuned in the future: interfaces can be included (Cheng & Shu, 2007) … room for progress
Interfaces in short!

Incompatible with high frequency solution as we request to have a smooth medium.

If an interface which have a smooth pattern, reset the problem by computing the asymptotic solution on the interface and restart a new problem (reflection or transmission at your own choice) using the solution as the initial solution: this is defined by the signature of the ray (PPSP ray for example).

If abrupt interface (no first derivative), go to geometrical theory of diffraction (Keller, 1962) or (Klem-Musatov, 1994) ...

Maybe too complex to manipulate for what we need in seismology and seismics?

Towards whispiring galleries (Babic, 1962; Thomson, 1995).
Interfaces in short!

Grab expertise from synthetic image software (Pixar)!

But you need to curve your rays!
Interfaces in short!

- Lagrangian approach: resume solutions at the interface and restart the ODE integration (Snell-Descartes law and paraxial Snell Descartes law) (Farra et al, 1989)

- Semi-lagrangian approach: procedure back to ODE (Rawlison & Sambdrige, 2003)

- PDE: rely on discontinuous finite element methods (Cheng & Shu, 2007) or boundary formulation
Conclusion Rays and Waves

- Geometrical optics: ODE versus PDE
  - Choose PDE when possible!

- ODE: tracing one (paraxial) ray is fast
  - Please trace paraxial rays as incremental cost

- Keep complexity low (seismic waves are finite frequency waves)
  - Do not drown yourself into the no-scale optical singularities

- Identification of rays is a key problem
  - PDE seems to solve it explicitly!
Perspectives Rays and Waves

- Fast sweeping method $O(N)$ for travel-times and for amplitudes
  - Amplitude equations have been designed

- Finite element methods put into the scene
  - Stencils are moving to higher orders and h-adaptivity

- Discontinuous Galerkin methods
  - This is the road to take for interface investigation in the frame of PDE.
Stop and move to the numerical exercise
http://seiscope.oca.eu

T H A N K  Y O U !
Geometrical optics models and numerical methods

\[ \Delta u + \omega^2 n(x)^2 u = 0 \]

- **Rays**
  \[ \frac{dx}{dt} = c^2 p, \quad \frac{dp}{dt} = -\frac{\nabla c}{c} \]
  - Ray tracing

- **Kinetic**
  \[ f_t + c^2 p \cdot \nabla_x f \]
  \[ -\frac{1}{c} \nabla c \cdot \nabla_p f = 0 \]
  - Wavefront methods
  - Moment methods
  - Phase space methods

- **Eikonal**
  \[ |\nabla \phi| = n(x) \]
  - Hamilton–Jacobi methods
Drawback: only first-arrival times!

Performing the forward problem and the adjoint problem allows one to estimate efficiently the gradient of the misfit function without rays.

Only the gradient and not the Jacobian matrix.

A VERY GOOD TOOL for First Arrival Time Tomography (FATT).

Four successive sweeps and one iteration (Taillardier et al, 2009).

An alternative to two tomographic techniques:
- Inversion through rays
- Inversion through times without ray tracing using simulated annealing (Asad et al, 1999)
Inverse problem

Jean Virieux

From Brossier (2013,2014)

Year 2015-2016

This could be the discovery of the century. Depending, of course, on how far down it goes...

Other inverse problems?
Outline of this partial overview

- Motivation and hints
- Travel time tomography (ABEL/HWB)
- Delayed travel time tomography
- Least-squares solution
- Quality control
- Applications
- Conclusion & Perspectives
Few hints
Least-square method

Sum of vertical distances between data points and expected y values from the unknown line $y=ax+b$ should be minimum: find a and b?

It is an inversion ....

From Excel

<table>
<thead>
<tr>
<th>DATA</th>
<th>MODEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>$y'$</td>
</tr>
<tr>
<td>2</td>
<td>2.42857</td>
</tr>
<tr>
<td>4</td>
<td>3.25714</td>
</tr>
<tr>
<td>6</td>
<td>4.08571</td>
</tr>
<tr>
<td>8</td>
<td>4.91429</td>
</tr>
<tr>
<td>10</td>
<td>5.74286</td>
</tr>
<tr>
<td>12</td>
<td>6.57143</td>
</tr>
<tr>
<td>14</td>
<td>7.4</td>
</tr>
</tbody>
</table>
Occam’s razor

- Do not use more complicated maths than the data deserves
- Approximate the least constrained quantity

Given: data (observed and modeled)
Assumed: wave propagation
Unknown: Earth structure

- Occam’s Razor: parcimonious principle

When you have many explanations for predicting exactly the same quantities and that there is no way to distinguish them, select the simplest one… until you end up with a contradiction.
Approximations can’t be avoided …

- Incomplete data
- Limited understanding
- Limited crunching capacity
  - Limited frequency range: translucent Earth
  - Limited dataset: direct access impossible (or difficult)
  - Data windowing (observables): extraction of robust information
  - Dimensionality (2D vs 3D …)
  - Structural complexity (multi-scale heterogeneities)
  - Physics approximation (elastic vs acoustic)

« An approximate solution to a real problem is better than an exact solution to an ideal problem » (Tarantola, 2007)
The sparse analysis cycle...

parsimonious modeling

Seismic wave propagation

Model space: 3D wavespeeds

robust pattern recognition

selecting informative data

Inverse modeling/imaging

sparsity constraints for inference

From Nissan-Meyer (Oxford)
How to reconstruct the velocity structure?

- **Forward problem (easy)**
  
  from a known velocity structure, it is possible to compute travel times, emergent distance and amplitudes.

- **Inverse problem (difficult)**
  
  from travel times (or similarly emergent distances), it is possible to deduce the velocity structure: this is the time tomography.

  even more difficult is the diffraction tomography related to the waveform and/or the preserved/true amplitude.
Tomographic approach

- Very general problem
  medicine; oceanography, climatology

- Difficult problem when unknown a priori medium (travel time tomography)

- Easier problem if a first medium could be constructed: perturbation techniques can be used for improving the reconstruction (delayed travel time tomography)
Travel Time tomography

- We must « invert » the travel time or the emergent distance for getting $z(u)$: we select the distance.

\[ X(p) = 2 \int_0^{z_p} \frac{p}{\sqrt{u^2(z) - p^2}} \, dz ; \quad \frac{X(p)}{2p} = \int_{u_0^2}^{p^2} \frac{dz}{\sqrt{u^2 - p^2}} \, du \]

- Abel problem (1826)

Determination of the shape of a hill from travel times of a ball launched at the bottom of the hill with various initial velocity and coming back at the initial position.
ABEL PROBLEM

A point of mass ONE and initial velocity \( v_0 \) reaches a maximal height \( x \) given by

\[
2v_0^2 = gx
\]

We shall take as the zero value for the potential energy: this gives us the following equations and its integration:

\[
\left( \frac{ds}{dt} \right)^2 = 2g(x - \xi) \quad ; \quad t(x) = \int_0^x \frac{ds}{d\xi} \frac{ds}{\sqrt{2g(x - \xi)}} d\xi
\]

We may transform it into the so-called Abel integral

\[
t(x) = \int_0^x \frac{f(\xi)}{\sqrt{x - \xi}} d\xi
\]

where \( t(x) \) is known and \( f(\xi) \) is the shape of the hill to be found: this is an integrale equation.
The exact inverse solution

We multiply and we integrate

\[ \int_0^\eta \frac{t(x)}{\sqrt{\eta - x}} \, dx = \int_0^\eta \frac{dx}{\sqrt{\eta - x}} \int_0^x \frac{f(\xi)}{\sqrt{x - \xi}} \, d\xi \]

We inverse the order of integration

\[ \int_0^\eta \frac{t(x)}{\sqrt{\eta - x}} \, dx = \int_0^\eta \frac{f(\xi) \, d\xi}{\sqrt{\eta - x}} \int_0^\xi \frac{dx}{\sqrt{\eta - x} \sqrt{x - \xi}} \]

We change variable of integration

\[ \int_0^\eta \frac{t(x)}{\sqrt{\eta - x}} \, dx = \pi \int_0^\eta f(\xi) \, d\xi \]

We differentiate and write it down the final expression

\[ \frac{d}{d\eta} \int_0^\eta \frac{t(x) \, dx}{\sqrt{\eta - x}} = \pi \cdot f(\eta) \]

\[ f(\xi) = \frac{1}{\pi} \frac{d}{d\xi} \int_0^\xi \frac{t(x) \, dx}{\sqrt{\xi - x}} \]
By changing variable $\xi$ in $a-\xi$ and $x$ in $a-x$, we get the standard formulae

We must have

- $t(x)$ should be continuous,
- $t(0)=0$
- $t(x)$ should have a finite derivative with a finite number of discontinuities.

The most restrictive assumption is the continuity of the function $t(x)$. 

$$
\begin{align*}
  t(x) &= \int_{a}^{x} \frac{f(\xi)}{\sqrt{\xi - x}} \, d\xi \\
  f(\xi) &= -\frac{1}{\pi} \frac{d}{d\xi} \int_{\xi}^{a} \frac{t(x) \, dx}{\sqrt{x - \xi}}
\end{align*}
$$
The solution HWB : HERGLOTZ-WIECHERT-BATEMAN

\[
\frac{X(p)}{2p} = \int_{u_0^2}^{p^2} \frac{dz / du^2}{\sqrt{u^2 - p^2}} du
\]

\[
z(v) = -\frac{1}{\pi} \int_{u_0^2}^{u^2} \frac{X(p) / 2p}{\sqrt{p^2 - u^2}} dp
\]

\[
z(v) = -\frac{1}{\pi} \int_{u_0}^{u} \frac{X(u)}{\sqrt{p^2 - u^2}} dp
\]

\[
z(v) = \frac{1}{\pi} \int_{0}^{X(u)} \frac{dX}{\cosh(p v)}
\]

From the direct solution, we can deduce the inverse solution

After few manipulations, we can move from the Cartesian expression towards the Spherical expression

We find \( r(v) \) as a value of \( r/v \)

In Cartesian frame

In spherical frame
Stratified medium

When considering discontinuities, ABEL/HWB method based on first arrival times has not an unique solution.

More over there will be an ambiguity when the velocity decreases (can only defined the velocity jump in this zone)

In fact, we avec an infinity of solutions when considering only direct and refracted waves.

We may find interfaces when considering all waves (including reflection waves)
Phase signature

- P : mantle P wave
- S : mantle S wave
- K : Outer core P wave
- I : Inner core P wave
- J : Inner core S wave
- c : Reflected P wave (outer core)
- i : Reflected P wave (inner core)
- m : number of reflections

Please, describe the wave with signature

PKP, PKKP, SKKKS=S3KS
Velocity structure with depth

- Velocity profile built without any a priori information.

An difficulty arises when the velocity decreases.
An initial model through the HWB method

- An initial model can be built
- The exact inverse formulae does not allow to introduce additional information,
  
  F. Press in 1968 has preferred the exhaustive exploration of possible profiles (5 millions !). The quality of the profile is appreciated using a misfit function as the sum of the square of delayed times as well as total volumic mass and inertial moments well constrained from celestial mechanics ...
- Exploration through grid search, Monte Carlo search, simulated annealing, genetic algorithm, tabou method, hant search …
The symmetrical radial EARTH
Time tomography

Only for one dimension … uptonow

What can we do in 2D, 3D and 4D?
A simple case: small perturbation

- Initial structure of velocity
- Search of small variation of velocity or slowness
- Linear approach
Example: Massif Central
Velocity variation at a depth of 200 km: good correlation with superficial structures.

Velocity variations at a depth of 1325 km: good correlation with the Geoid.

Courtesy of W. Spakman
Delayed Travel-time tomography

Finding the slowness \( u(x,y,z) \) from \( t(s,r) \) is a difficult problem: only techniques for one variable \( u(z) \) (Abel)!

Consider small perturbations \( \delta s(x,y,z) \) of the slowness field \( s_0(x,y,z) \)

\[
t(s,r) = \int_s^r s(x,y,z) dl = \int_s^r s_0(x,y,z) dl + \int_s^r \delta s(x,y,z) dl
\]

\[
t(s,r) = \int_{s_0}^{r_0} s_0(x,y,z) dl + \int_{s_0}^{r_0} \delta s(x,y,z) dl
\]

\[
t(s,r) - t_0(s,r) = \int_{s_0}^{r_0} \delta s(x,y,z) dl
\]

\[
\delta t(s,r) = \int_{s_0}^{r_0} \delta s(x,y,z) dl
\]

LINEARIZED PROBLEM \( \delta t(d) = J(d,m) \delta s(m) \)

from the model domain to the data domain
Discretization of the slowness perturbation

The velocity perturbation field (or the slowness field) $\delta s(x, y, z)$ can be described into a meshed cube regularly spaced in the three directions.

For each node, we specify a value $\delta s_{i,j,k}$. The interpolation will be performed with functions as step functions. For each grid point $(i,j,k)$, shape functions $h_{i,j,k} = 1$ for $i,j,k$, and zero for other indices.

$$\delta s(x, y, z) = \sum_{cube} \delta s_{i,j,k} h_{i,j,k}$$

Nodal approach

Other shape functions are possible with two end members (nodal versus global):

- fourier functions ($\cos, \sin$), chebychev, spline … and so on
Discrete Model Space

\[ \delta t(s, r) = \int_{ray_0}^{\text{cube}} \sum_{i,j,k} \delta s_{i,j,k} h_{i,j,k} \, dl = \sum_{i,j,k} \delta s_{i,j,k} \int_{ray_0}^{\text{cube}} h_{i,j,k} \, dl \]

Discretization of the medium fats the ray

\[ \delta t(s, r) = \sum_{i,j,k} \delta s_{i,j,k} \Delta l_{i,j,k} = \sum_{i,j,k} \frac{\partial t}{\partial s_{i,j,k}} \delta s_{i,j,k} \]

\[ \delta t(s, r) = \sum_{i,j,k} J_{i,j,k} \delta s_{i,j,k} \]

\[ \delta t(d) = J(d, m) \, \delta s(m) \]

To be solved in least squares sense

Sensitivity matrix J is a sparse matrix

also named Fréchet derivative or Jacobian matrix …

24/02/2016

HR seismig imaging
Linear algebra

$$\delta t = J \delta u$$

$$d = Gm$$

$$b = Ax$$
The rectangular system can be recast into a square system (sometimes called normal equations).

Solving this square linear system gives the so-called least-squares solution.

Another interesting solution

The system is both under-determined and over-determined depending on the considered zone (and the number of rays going through.)
LEAST SQUARES METHOD

\[ E(m) = (d - G_0 m)^t (d - G_0 m) \]  
\[ \frac{\partial E(m)}{\partial m} = 0 \]  
\[ G_0^t G_0 m = G_0^t d \]  
\[ m_{est} = \left( G_0^t G_0 \right)^{-1} G_0^t d \]  

L₂ norm  
G₀ is a N by M matrix  
\( \left[ G_0^t G_0 \right]^{-1} \) is a M by M matrix  

Least-squares estimation

Operator \( \left[ G_0^t G_0 \right]^{-1} G_0^t \) on data will derive a new model: this is called

the generalized inverse \( G_0^g \)

Under-determination \( M > N \)  
Over-determination \( N > M \)

Mixed-determination – seismic tomography
SVD analysis for stability and uniqueness

SVD decomposition:

\[ G_0 = U \Lambda V^t \]

- \( U \) : \((N \times N)\) orthogonal, \( U^t = U^{-1} \)
- \( V \) : \((M \times M)\) orthogonal, \( V^t = V^{-1} \)
- \( \Lambda \) : \((N \times M)\) diagonal matrix

Null space for \( \lambda_i = 0 \)

\[ U^t U = I \text{ and } V^t V = I \] (not the inverse !)

\[ V = [V_p | V_0] \]
\[ U = [U_p | U_0] \]

\( V_p \) and \( V_0 \) determine the uniqueness while \( U_p \) and \( U_0 \) determine the existence of the solution

\[ G_0 = U_p \Lambda_p V_p^t \]
\[ G_0^{-1} = V_p \Lambda_p^{-1} U_p^t \]

Up and \( V_p \) have now inverses!

24/02/2016
The solution is

$$m_{est} = G_0^{-1} d = G_0^{-1} (G_0 m) = V_p \Lambda_p^{-1} U_p^t U_p \Lambda_p V_p^t m = [VV_p^t]m = Rm$$

where $$R = [V_p \ V_p^t]$$  Model resolution matrice : if $$V_0 = 0$$ then $$R=VV^t=I$$

$$d_{est} = G_0 m_{est} = [U_p U_p^t]d = Nd$$

where $$N = [U_p U_p^t]d$$  Data resolution matrice  : if $$U_0 = 0$$ then $$N=UU^t=I$$

importance matrice

Goodness of resolution

$$\text{SPREAD}(R) = \left\| R - I \right\|^2$$

Good tools for quality estimation

$$\text{SPREAD}(N) = \left\| N - I \right\|^2$$

Spreading functions

24/02/2016

HR seismig imaging
Updating slowness perturbation values from time residuals

Formally one can write

\[ \delta u = G_0^{-1} \delta t \rightarrow m = G_0^{-1} d \]

with the forward problem

\[ \delta t = G_0 \delta u \rightarrow d = G_0 m \]

Existence, Uniqueness, Stability, Robustness

Discretisation, Identifiability, Small errors propagates, Outliers effects

NON-UNIQUENESS & NON-STABILITY : ILL-POSED PROBLEM

REGULARISATION : ILL-POSED -> WELL-POSED
Linearized Inverse Problem

Misfit function

\[ C(s) = \frac{1}{2} \delta t^t \delta t \]

- First loop over models: for current model \( s_k \) (iteration \( k \))
  - Forward problem
    - A. Solve eikonal equation
    - B. Compute rays and synthetic travel-times
  - Build Fréchet derivative matrix \( J_k \) and delayed times \( \delta t_k \)

- Second loop over linear system: iteratively solve \( J_k \delta s_k = \delta t_k \), ie \( J_k^t J_k \delta s_k = \delta t_k \) using conjugate gradient (LSQR, for example)

Update the model \( s_{k+1} = s_k + \delta s_k \)

No explicit evaluation of the Hessian \( J_k^t J_k \): only products « \( J \delta s \) » and « \( J^t \delta t \) » are required in LSQR algorithm

Paige & Sanders (1982)
Linearized Inverse Problem

Misfit function

\[ C(s) = \frac{1}{2} \delta t^t \delta t \]

- First loop over models: for current model \( s_k \) (iteration \( k \))
  - Forward problem
    A. Solve eikonal equation
    B. Compute rays and synthetic travel-times
  - Build Fréchet derivative matrix \( J_k \) and delayed times \( \delta t_k \)

- Second loop over linear system: iteratively solve \( J_k \delta s_k = \delta t_k \), ie \( J_k^t J_k \delta s_k = \delta t_k \) using conjugate gradient (LSQR, for example)

Update the model \( s_{k+1} = s_k + \delta s_k \)

Complexity: \( \mathcal{O}(N_{src} \times N_{rec}) \) for forward/adjoint formulation and
\[ \mathcal{O}(N_m \times N_{src} \times N_{rec} \times k) \] for gradient storage (\( k \): sparsity)
Misfit function

\[ \mathcal{C}(s) = \frac{1}{2} \delta t^t \delta t \]

- First loop over models: for current model \( s_k \) (iteration \( k \))
  - Forward problem
    - A. Solve eikonal equation
    - B. Compute synthetic travel-times and delayed times \( \delta t_k \)
  - Adjoint problem
    - A. Solve adjoint equation
    - B. Build gradient \( \gamma_k \)

- Descent method (one-step gradient or conjugate gradient or l-BFGS) \( \delta s_k \)

Update the model \( s_{k+1} = s_k + \alpha_k \delta s_k \)

(Sei & Symes, 1994; Leung & Qian, 2006; Taillandier et al, 2009)

No explicit evaluation of the Hessian
I-BFGS provides an evaluation of the \( H^{-1} \) from previous stored gradients
Linearized Inverse Problem

Misfit function

\[ \mathfrak{C}(s) = \frac{1}{2} \delta t^t \delta t \]

- First loop over models: for current model \( s_k \) (iteration \( k \))
  - Forward problem
    - A. Solve eikonal equation
    - B. Compute synthetic travel-times and delayed times \( \delta t_k \)
  - Adjoint problem
    - A. Solve adjoint equation
    - B. Build gradient \( \gamma_k \)

- Descent method (one-step gradient or conjugate gradient or l-BFGS)
  \[ \delta s_k \]

Update the model \( s_{k+1} = s_k + \alpha_k \delta s_k \)
(Sei & Symes, 1994; Leung & Qian, 2006; Taillandier et al, 2009)

Complexity: \( \mathcal{O}(N_{src}) \) for forward/adjoint formulation and \( \mathcal{O}(Nm) \) for gradient
_linearized Inverse Problem

Misfit function

\[ C(s) = \frac{1}{2} \delta t^t \delta t \]

Two-loop procedure

- **First loop over models**: for current model \( s_k \) (iteration \( k \))
  - **Forward problem**
    - A. Solve eikonal equation
    - B. Compute synthetic travel-times and delayed times \( \delta t_k \)
  - **Adjoint problem**
    - A. Solve adjoint equation
    - B. Build gradient \( \gamma_k \)

- **Second loop over linear system**: \( H_k \delta s_k = -\gamma_k \)
  - Truncated Newton method based on iterative conjugate gradient
    (matrix free approach for product \( H_k \nu \), thanks to second-order adjoint formulation)

Update the model \( s_{k+1} = s_k + \alpha_k \delta s_k \)

(Métivier et al, 2013)

Full Hessian impact \( H = J^t J + \frac{\partial J}{\partial s} \delta t \) is evaluated
Maximum Likelihood method

One assume a \textit{gaussian distribution} of data

Data distribution could be written

\[ p(d) \propto \exp \left[ -\frac{1}{2} (d - \hat{G}_0 m)^t C_d^{-1} (d - \hat{G}_0 m) \right] \]

where \( \hat{G}_0 m_{est} \) is the data mean and \( C_d \) is the data covariance matrice.

This method is very similar to the least squares method where we define the following misfit function \( E_1 \).

\[ E(m) = (d - \hat{G}_0 m)^t (d - \hat{G}_0 m) \rightarrow E_1(m) = (d - \hat{G}_0 m)^t C_d^{-1} (d - \hat{G}_0 m) \]

Even without knowing the matrice \( C_d \), we may consider data weight \( W_d \) through the misfit function

\[ E_2(m) = (d - \hat{G}_0 m)^t W_d (d - \hat{G}_0 m) \]

The inferred model distribution will be gaussian with the main difficulty of estimating the posterior model covariance \( C_m \) connected with the curvature of the misfit function \( E_2 \).
PRIOR INFORMATION

Hard bounds \( A < m_i < B \)  

Seismic velocity should be positive

Prior model

\[
E_3(m) = (d - G_0 m)^t (d - G_0 m) + \varepsilon (m - m_p)^t (m - m_p)
\]

The parameter \( \varepsilon \) is the **damping parameter** controlling the importance of the model \( m_p \)

Gaussian distribution

\[
E_4(m) = (d - G_0 m)^t C_d^{-1} (d - G_0 m) + (m - m_p)^t C_m^{-1} (m - m_p)
\]

\[
G_0^g = \left[ G_0^t C_d^{-1} G_0 + C_m^{-1} \right]^{-1} G_0^t
\]

Model smoothness

\[
E_5(m) = (d - G_0 m)^t W_d (d - G_0 m) + (m - m_p)^t W_m (m - m_p)
\]

with \( W_d \) data weighting and \( W_m \) model weighting

Penalty approach

add additional relations between model parameters (new lines)
L curve

Estimating the parameter $\varepsilon$ is always difficult

We can test different values which provide a L curve

If a knee appears, this could be an adequate value … although this knee is often difficult to identify in seismic tomography.

Other techniques are available as LASSO approach but they are intensive and sensitive to noise.
UNCERTAINTY ESTIMATION

Least squares solution

\[ m_{est} = \left[ G_0^t G_0 \right]^{-1} G_0^t d = G_0^g d \]

\[
\begin{bmatrix}
\text{cov} & m_{est}
\end{bmatrix} = G_0^g \begin{bmatrix}
\text{cov} & d
\end{bmatrix} G_0^{gt} = G_0^g C_d G_0^{gt}
\]

\[ C_d = \sigma_d^2 I \]

Uncorrelated data

\[
\begin{bmatrix}
\text{cov} & m_{est}
\end{bmatrix} = \sigma_d^2 \left[ G_0^t G_0 \right]^{-1}
\]

\[
\begin{bmatrix}
\text{cov} & m_{est}
\end{bmatrix} = \sigma_d^2 \left[ \frac{1}{2} \frac{\partial^2 E}{\partial m^2} \right]^{-1}
\]

Model covariance:

- uncertainty in the data
- curvature of the error function

Sampling the misfit function around the estimated model:

often this has to be done numerically
A posteriori model covariance matrice

If one can decompose this matrice

\[ G_0^t C_d^{-1} G_0 + C_m^{-1} = USU^t \]

S diagonal matrice  eigenvalues
U orthogonal matrice eigenvectors

Error ellipsoidal could be estimated

WARNING : formal estimation related to the gaussian distribution hypothesis
A priori & A posteriori information

What is the meaning of the « final » model we provide ?

acceptable
Steepest descent methods

\[ E(m_{k+1}) < E(m_k) \]

Gradient method

\[ E(m_k - t\nabla E(m_k)) < E(m_k) \]

\[ d = -\nabla E(m_k) \]

Conjugate gradient

\[ d_k = \begin{cases} -\nabla E_0 \\ -\nabla E_k + \beta_k d_{k-1} \end{cases} \]

Newton

\[ d = -\nabla^2 E(m_k)^{-1} \nabla E(m_k) \]

Quasi-Newton

\[ \nabla^2 E(m_k) \approx D_k \]

Gauss-Newton is Quasi-Newton for L^2 norm
Tomographic descent

Minimisation of this vector

\[
\frac{1}{2} \left\| \begin{bmatrix}
C_d^{-1/2} d \\
- C_m^{-1/2} m_p
\end{bmatrix} - \begin{bmatrix}
C_d^{-1/2} g(m) \\
- C_m^{-1/2} m
\end{bmatrix} \right\|^2
\]

If one computes

\[
A_k = \begin{bmatrix}
- C_d^{-1/2} G_k \\
C_m^{-1/2}
\end{bmatrix}
\]

then

\[
(A_k^T A_k) \delta m = A_k^T \begin{bmatrix}
C_d^{-1/2} (g(m_k) - d) \\
C_m^{-1/2} (m_0 - m_k)
\end{bmatrix}
\]

Gaussian error distribution of data and of a posteriori model

Easy implementation once \( G_k \) has been computed

Extension using Sech transformation (reducing outliers effects while keeping \( L^2 \) norm simplicity)
The LSQR method is a conjugate gradient method developed by Paige & Saunders.

Good numerical behavior for ill-conditioned matrices.

When compared to an SVD exact solution, LSQR gives main components of the solution while SVD requires the entire set of eigenvectors.

Fast convergence and minimal norm solution (zero components in the null space if any).
Flow chart

\[ d_{\text{obs}} \] collected data
\[ m \] starting model

\[ d_{\text{syn}} = g(m) \] true ray tracing

\[ \Delta d = d_{\text{obs}} - d_{\text{syn}} \] data residual

\[ G_0 = \frac{\partial g}{\partial m} \] sensitivity matrice

\[ \Delta m = G_0^{-1} \Delta d \] model update

\[ m = m + \Delta m \] new model

\[ \text{small model variation or small errors exit} \]

Calculate \[ \frac{\partial^2 E}{\partial m^2} \] for formal uncertainty estimation

24/02/2016
Sampling a posteriori distribution

Resolution estimation: spike test
Boot-Strapping
Jack-knifing
Natural Neighboring
Monte-Carlo
Sampling a posteriori distribution

Uncertainty estimation for P and S velocities using boot-strapping techniques

24/02/2016 HR seismig imaging
Ray approximation

\[ \delta T(s, r) \approx \int u(x(l))dl \approx \iiint u(x)\delta(x - x(l))K(s, r, x)dv \]

\[ K(x, y, z) \text{ is the sensitivity kernel} \]

Line

Volume
Ray approximation reasonable approximation

Gradient $J^t \delta t$

Same gradient once estimated on the model discretization grid
Delayed **Travel-time** Tomography

\[
\delta T(s, r) = \int u(x(l))dl = \iiint u(x)\delta(x - x(l))K(s, r, x)dv
\]

Still DTT provides impressive images while we do believe that DFT would provide better images in the future, thanks to finer discretization coming with the densification of the available data.

(Li & van der Hilst, 2010)

\[
\delta T(s, r) = \iiint u(x) K(s, r, x)dv
\]
Delayed Fresnel Tomography

\[ \delta T(s, r) = \iiint u(x) K(s, r, x) dv \]

No amplitude measurements
Only phase information
Valid also for dispersive waves
Improved resolution?
Windows definition

(LTA/STA; Maggi et al, 2009)

(Wavelet Freq/time; Lee and Chen, 2013)
Phases measurements

(Zaroli et al, 2010)
Workflow for DFT

- Taking seismograms
- Windowing phases
- Picking phases (cross-correlation)
- Definition of a misfit function (L2)
- Delayed Fresnel Tomography
- Model perturbation (velocity)
Phase differences

We do not pick phases as for DFT:

(Fichtner et al, 2009)
Full Phases Inversion

(Fichtner et al, 2009)

Do we improve the resolution compared to DFT?

If equivalent, very expensive; if not, smooth transition to FWI
Workflow for FPI

- Taking seismograms
- Windowing phases
- Definition of a misfit function (time differences)
- Full phases inversion
- Model perturbation (velocity)
Applications
Corinth Gulf

An extension zone where there is a deep drilling project.

How this rift is opening?

What are the physical mechanisms of extension (fractures, fluides, isostatic equilibrium)

Work of Diana Latorre and of Vadim Monteiller

24/02/2016
Seismic experiment 1991 (and one in 2001)
MEDIUM 1D: HWB AND RANDOM SELECTION
Velocity structure image

Horizontal sections

Basin structure
Fast variation
Vp/Vs ratio: fluid existence?

Recovered parameters might have different interpretation and the ratio Vp/Vs has a strong relation with the presence of fluids or the relation Vp*Vs may be related to porosity.
Other methods of exploration

- Grid search
- Monte-Carlo (ponctual or continuous)
- Genetic algorithm
- Simulated annealing and co
- Tabou method
- Natural Neighboring method
Conclusion FATT

- Selection of an enough fine grid
- Selection of the a priori model information
- Selection of an initial model
- FMM and BRT for 2PT-RT
- Time and derivatives estimation
- LSQR inversion
- Update the model
- Uncertainty analysis (Lanzos or numerical)
Dramatic increase of acquisition density both in the academic world (increasing density) and in the industrial world (continuous recording).
Archive Size

As of January 1, 2011

Tenabytes

- EarthScope
- PASSCAL
- Engineering
- US Regional
- Other
- JSP
- FDSN
- GSN
THANK YOU!

Many figures have come from people I have worked with:

many thanks to them!
END HERE
THE \( C_m^{-1/2} \) MATRICE

The matrix \( C_m \) has a band diagonal shape
- \( \sigma \) is the standard error (same for all nodes)
- \( \lambda \) is the correlation length

\[
C_{ij} = \sigma^2 \exp\left(-\frac{|x_i - x_j|}{\lambda}\right)
\]

\( n = n_x n_y n_z = 10^4 \) \( C_m = U S U^t \)

(Lanzos decomposition)

\[
C_m^{-1/2} = US^{-1/2} U^t
\]
Analysis of coefficients

Values are only related to $\sigma$ and $\lambda$

$$n_0 C_m^{-1/2} \rightarrow n \tilde{C}_m^{-1/2}$$

Typical sizes 200x200x50
deduced from 20x20x5 (few minutes)

Strategy of libraries of $C_m^{-1/2}$ for various $\lambda$ and $\sigma=1$

Other coefficients could be deduced

R: $C_m^{-1/2}$ sparse matrice
An example

Same numerical grid for all simulations (either 100x100 or 400x400)

Same results at the limit of numerical precision related to the estimation of the sensitivity matrix

$$\sigma_v = 100 \, \text{km/s}$$
$$\sigma_x = 100 \, \text{km}$$
$$\sigma_t = 100 \, \text{s}$$

Ray imprints

$$\lambda = 0.1$$

$$\lambda = 0.8$$
Illustration of selection \{\lambda, \sigma_v\}

- \lambda = 3 \text{ km}, \sigma_v = 3 \text{ km/s}
- \lambda = 5 \text{ km}, \sigma_v = 3 \text{ km/s}

Error function analysis will give us optimal values of a priori standard error and correlation length (2D analysis).

\[ \lambda = 5 \text{ km and } \sigma_v = 3 \text{ km/s} \]
What is the meaning of the « final » model we provide ?
The probability density in the model space is related to the data space by the equation:

\[ \rho(d, m) = \rho_D(d) \rho_M(m) \]
Different informations

PROBLÈME DIRECT

INFORMATION A PRIORI

LOI A POSTERIORI

LOI MARGINALE
Perspectives

- A priori model covariance: location dependent wavelet decomposition allows local analysis (work of Matthieu Delost-Geoazur)
- Fresnel tomography: introduction of the first Fresnel zone influence in the forward problem (work of Stéphanie Gautier-Montpellier)
Kirchhoff approximation

1) Representation theorem
2) Kirchhoff summation
3) Reciprocity
source  

recepteur  

contribution speculaire  

interface  

Refl.(s)  

s
$V_{s1}^s = 2.5 \frac{km}{s}$  \hspace{1cm} $\rho_1 = 2.6 \frac{kg}{dm^3}$

$V_{s2}^s = 4. \frac{km}{s}$  \hspace{1cm} $\rho_2 = 3.1 \frac{kg}{dm^3}$
SISMOGRAMME CONVOLUE PAR UN SIGNAL SOURCE.
Born approximation

1) Single scattering approximation
2) Surface approximation
source

récepteurs

distance (m)

profondeur (m)

point diffractant

distance source-récepteurs (m)

temps (s)

signal source

amplitude

temps (s)
Forgues et al., 1996
http://seiscope2.osug.fr

THANK YOU!

It's not much...
but it's all I have