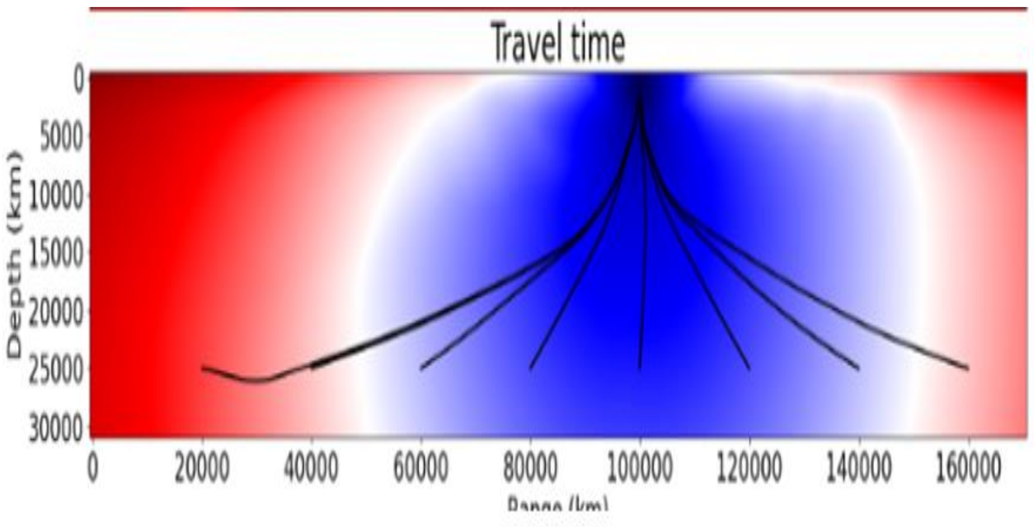
First-arrival traveltime tomography

First-break traveltime tomography: where are we?
Illustration on 30 years Western-Alps database

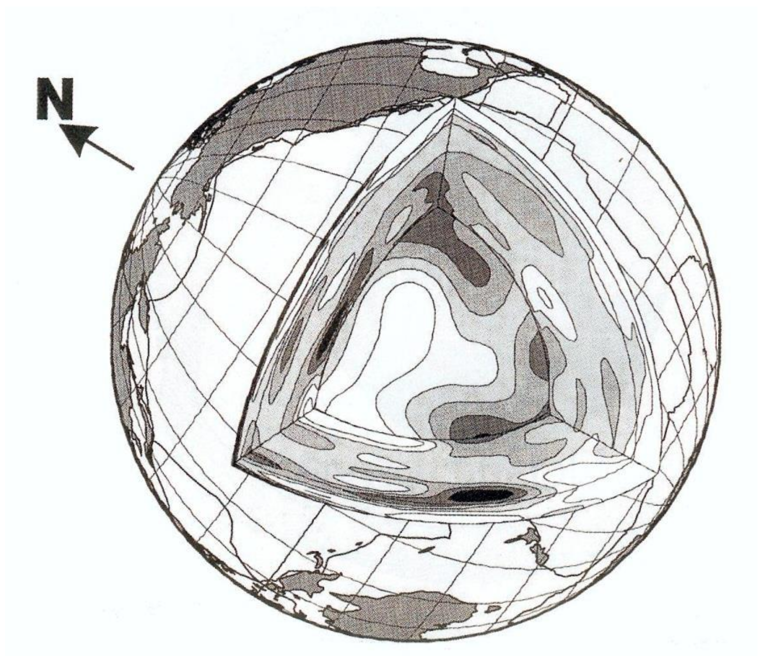
By Jean Virieux, Pr. Emeritus
Univ. Grenoble Alpes, ISTerre
France

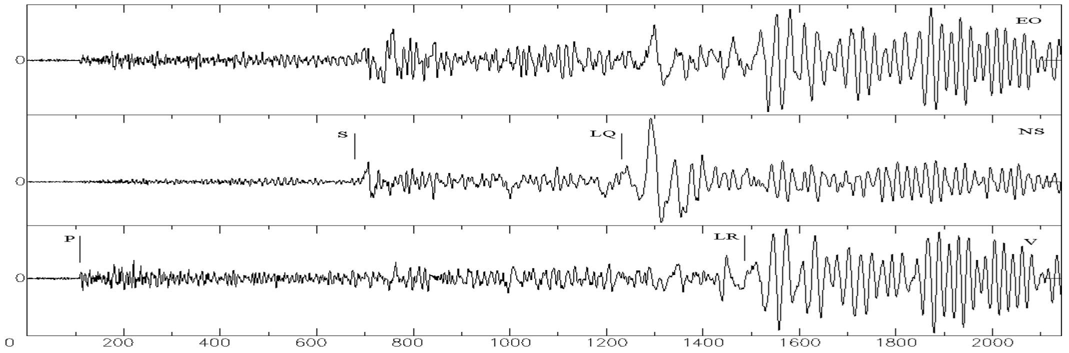


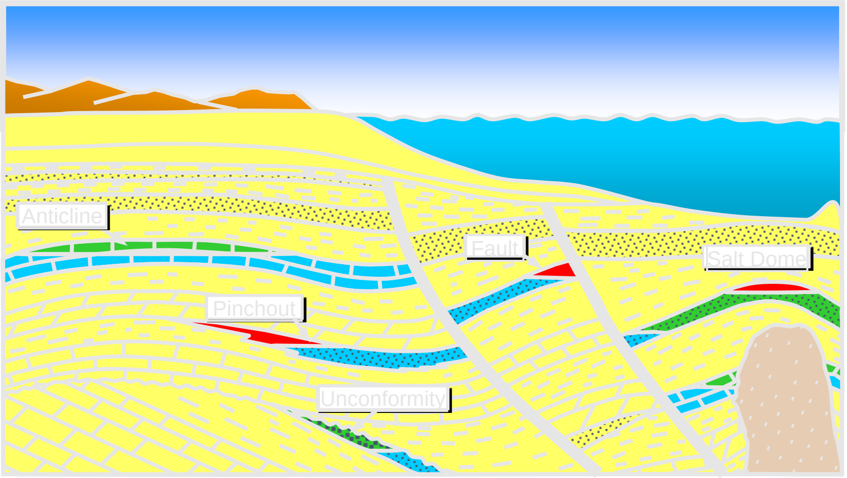
Seiscope consortium

AkerBP BR PETROBRAS Chevron CGG ExxonMobil JGI Statoil Schlumberger Shell Sinopec Total

Waves: phases and vibrations

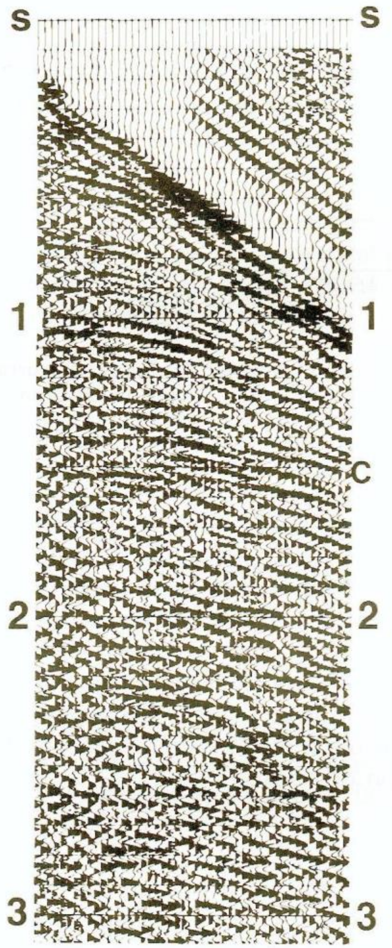






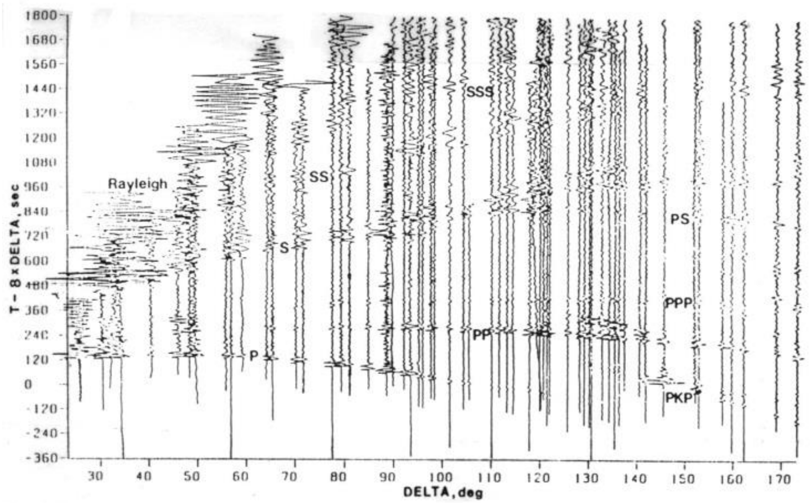
American Petroleum Institute, 1986

- Source time (rupture velocity)
  - from 0.1 sec to 100 sec
- Wave time
  - from seconds to hours
- Window time
  - from few seconds to days

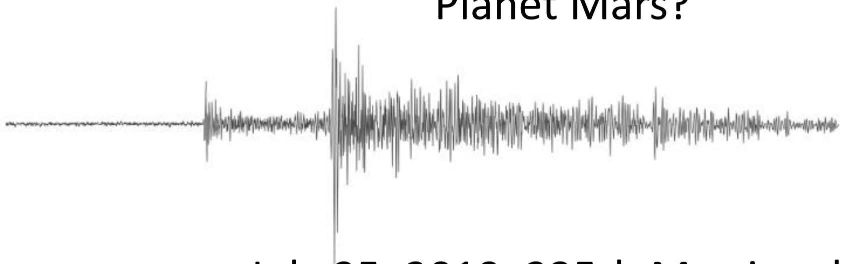


Seismograms

Records of a far-distance earthquake (Müller and King, 1976)

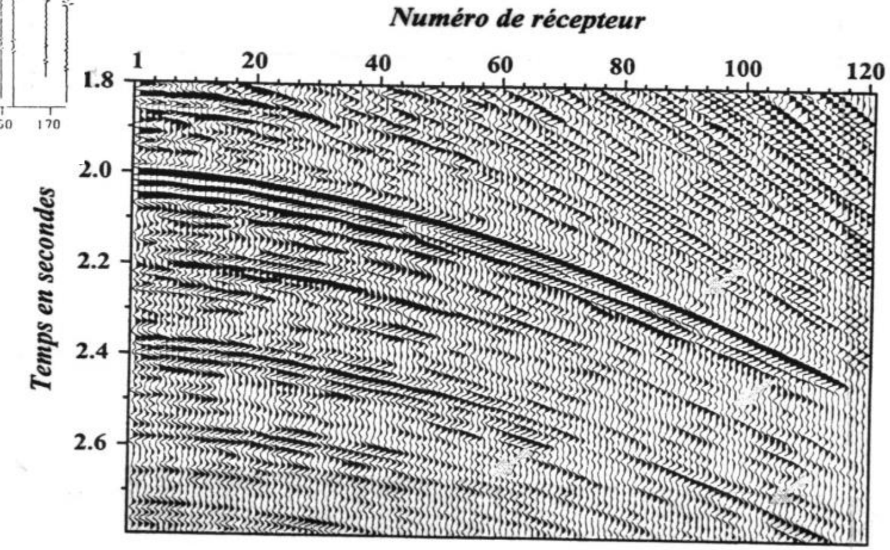


Planet Mars?

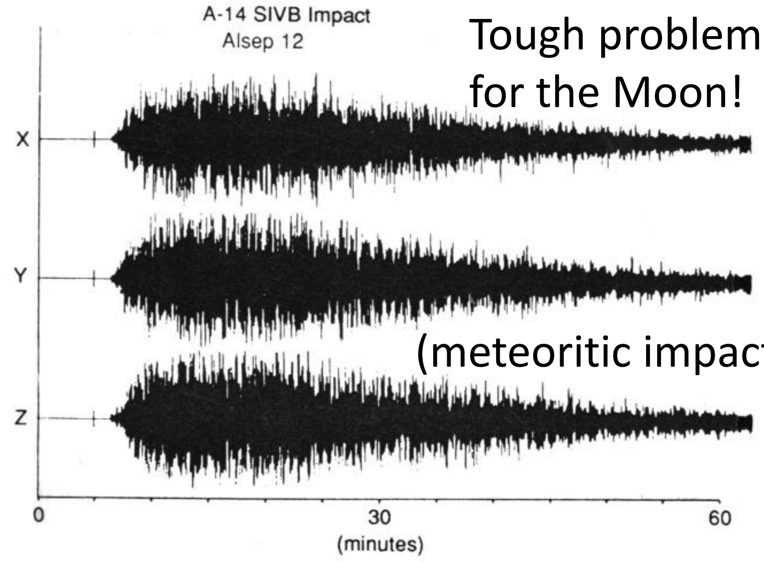


July 25, 2019, 235th Martian day

Traces from an oil reservoir (Thierry, 1997)

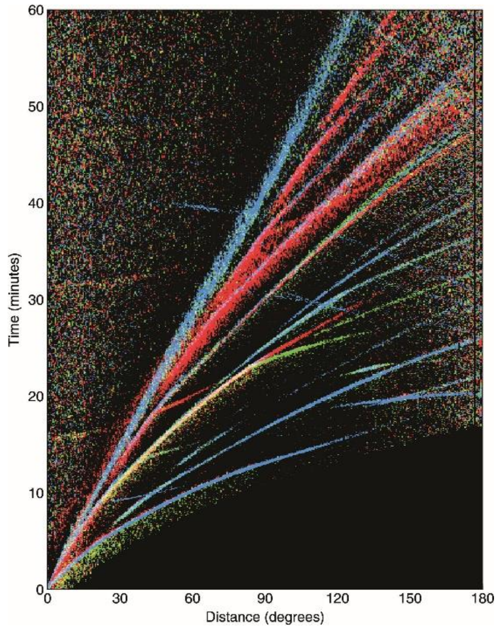


Records on the Moon (Latham et al., 1971)



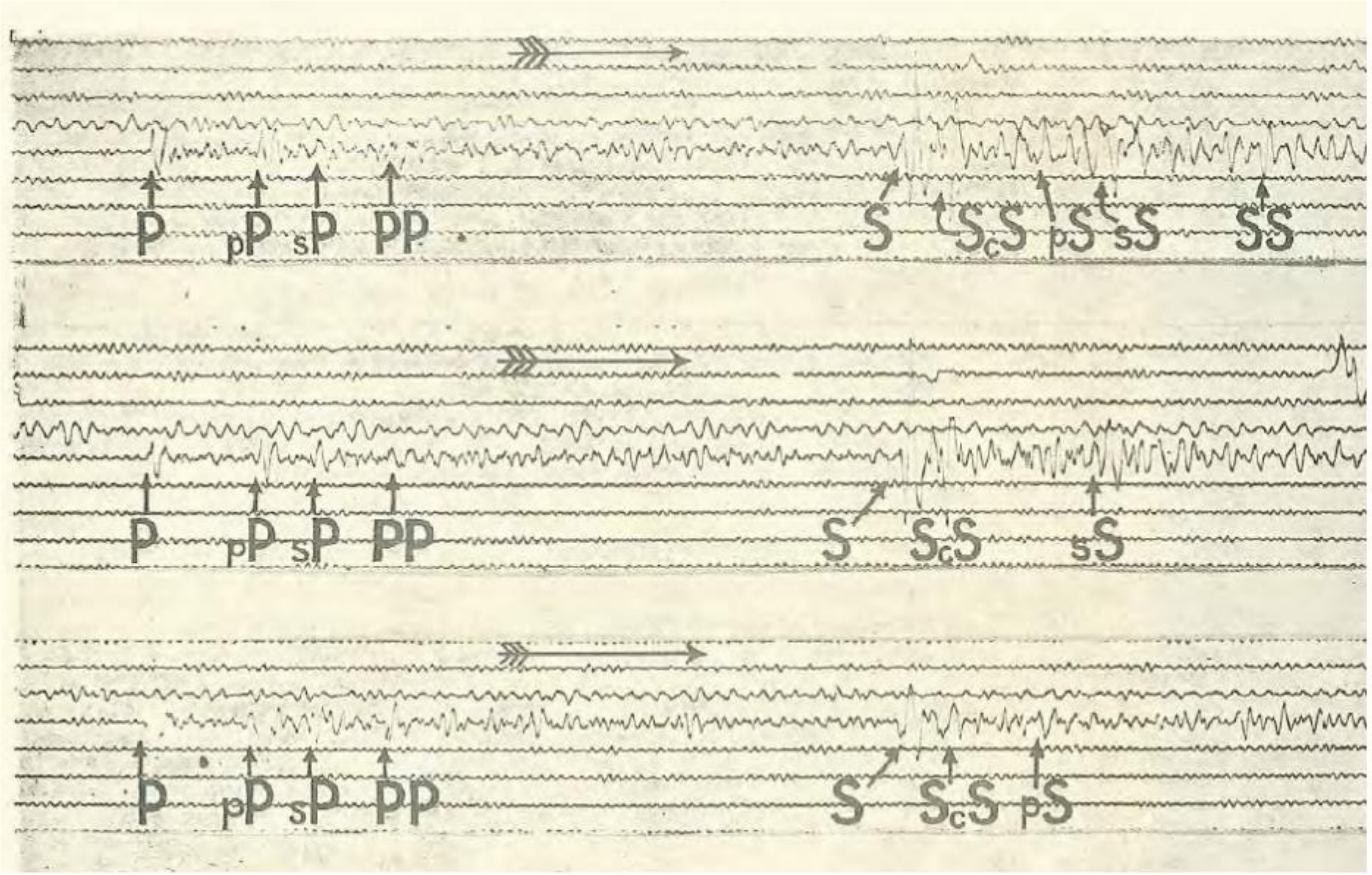
Tough problem for the Moon!

(meteoritic impact)



Impressive data mining over centuries: phase picking (ISC)

Onset time (date)



From Jeffreys, "The Earth"

Phase or time : a very robust observable to be extracted from seismograms

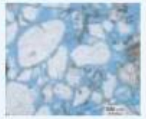
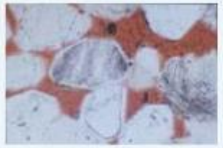
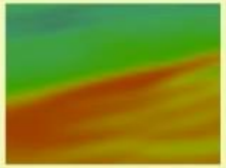
Outline on first-arrival traveltime tomography

- Images at very different scales
- Waves and Phases: various concepts
- Few points on first-break ray-based tomography
- Illustration on 30-years Western Alps tomography
- First-break eikonal-based tomography
- First-break wave-equation-based tomography
- Hypocenter-velocity joint inversion
- Conclusion

Expected resolution of seismic images

Translucent Earth

Unaccessible target…
Intrinsic remote sensing limitation…
Seismic finite frequency & attenuation…

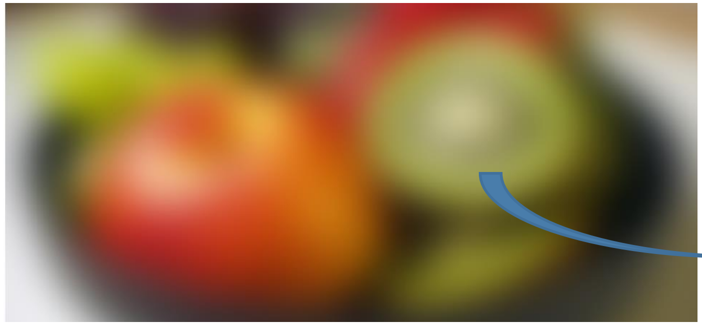
  µm	**Micro-scale**	Physical parameters: mineral composition, gas, liquid …
	upscaling ↓ / downscaling ↑	
 < or ~ m	**Meso-scale**	Inference parameters: one solid skeleton and one fluid. Gassmann rheology Porosity, saturation, tortuosity, consolidation parameter …
 > 10*m	**Macro-scale** (Investigated by FWI)	Model parameters: velocities, attenuation, anisotropy, density for seismic waves

Inspired from Depuy's PhD (2011)

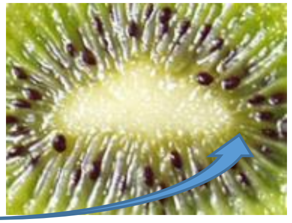
Real medium

Other information
Geology
Rock physics
Remote sensing

How far could we go with phases?



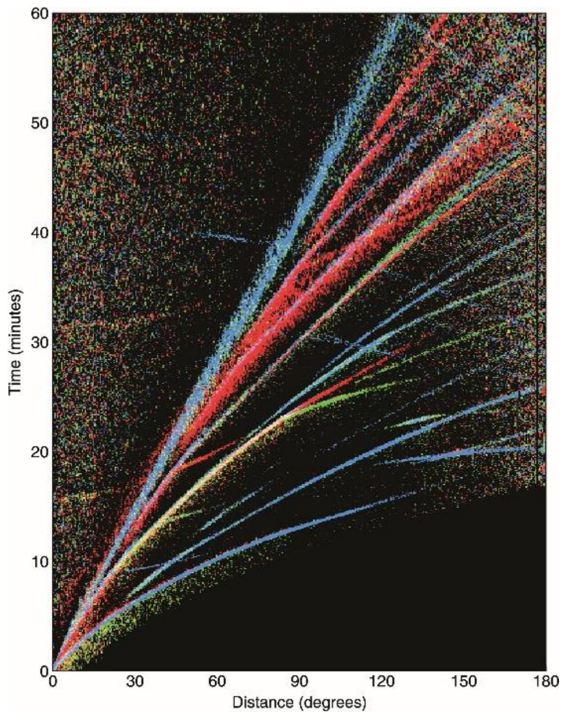
Interpretation?



Seismic imaging

Inspired from Romanowicz's lecture

Seismic images from time/phase picks



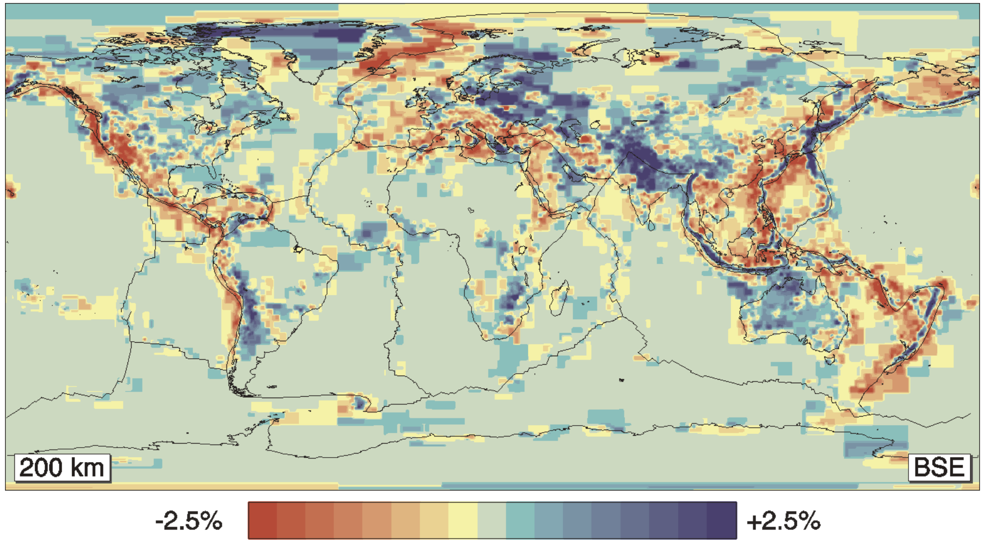
Accumulation of picks over centuries …

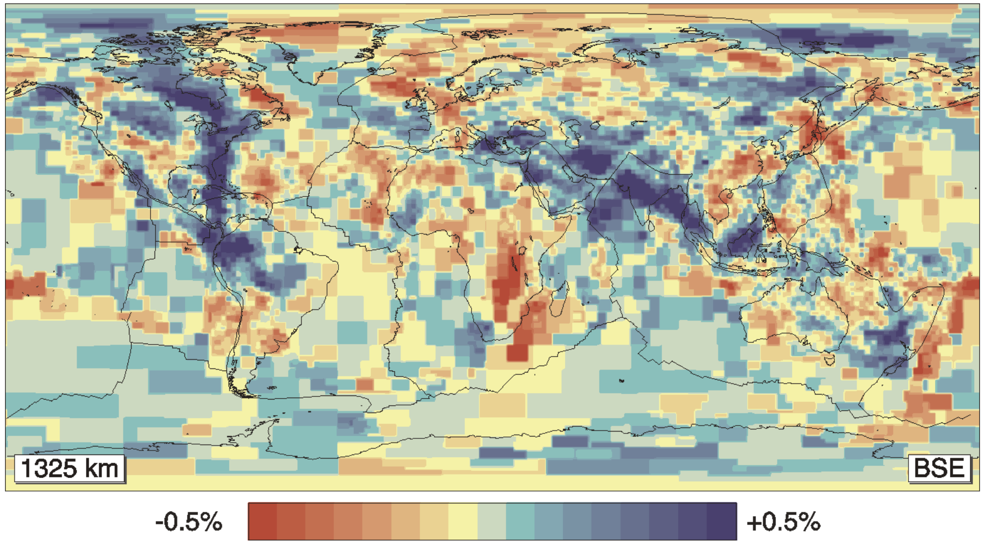
- Velocity variation at a depth of 200 km : good correlation with superficial structures.

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

(Courtesy of W. Spakman)

Earthquake positions are fixed





Global scale

Seismic images at all scale variations

Impressive scale variation

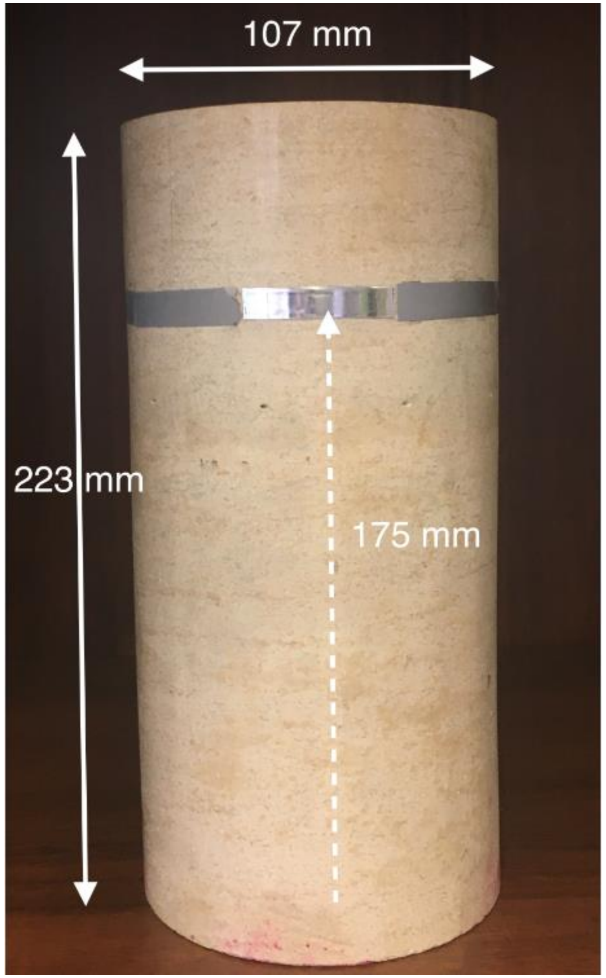
- Global scale
- Lithospheric/continental scale
- Upper crustal scale
- Near-surface scale
- Laboratory scale

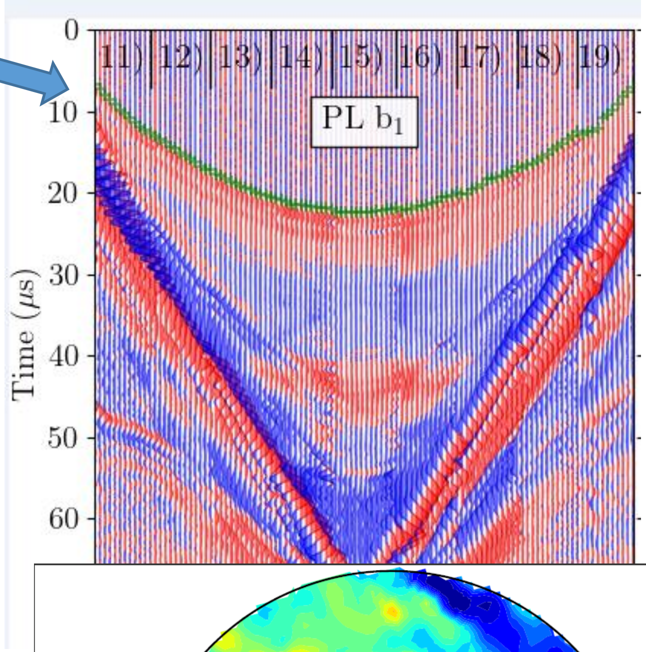
Physical upscaling of first-arrival times: crossing scale variation

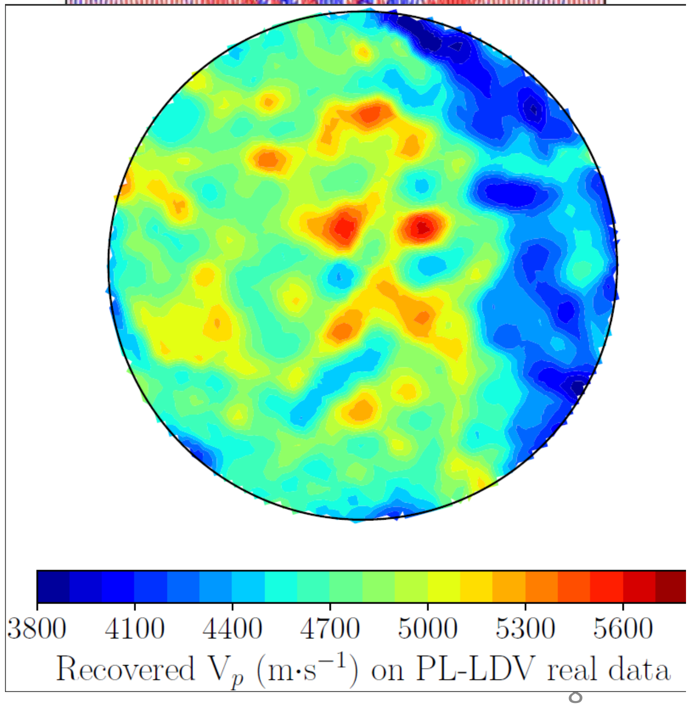
First-arrival time tomography
Results may cross scale variation …

Only first-arrival picks

Carbonate sample

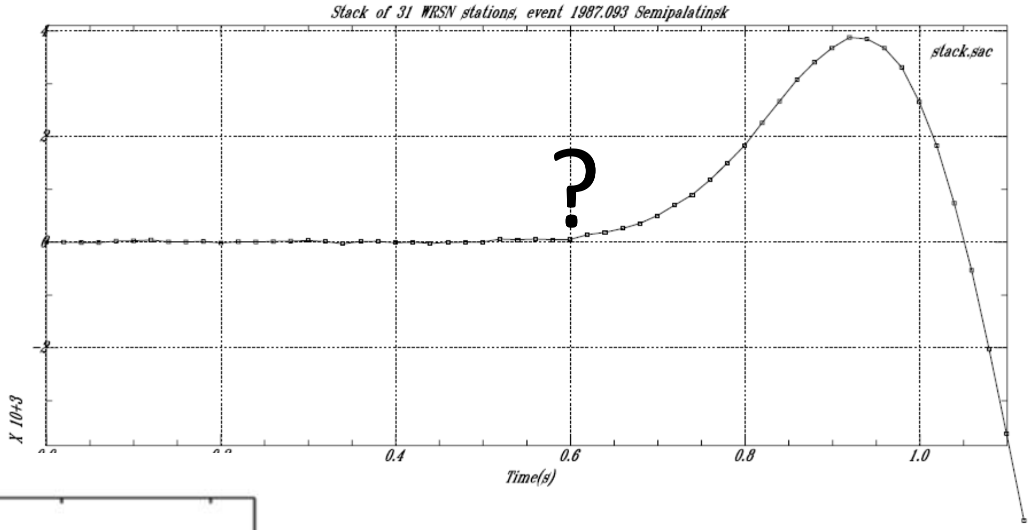






Shen et al., (in preparation)

Picking onset time



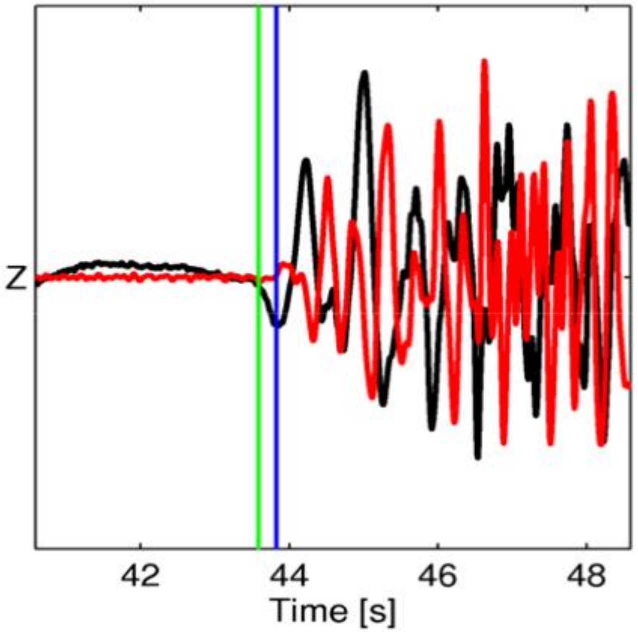
From Nolet (2010)

Onset time: picking and association (Pn, Pg …)

Agnostic to frequency content

Diffraction with wavefront healing

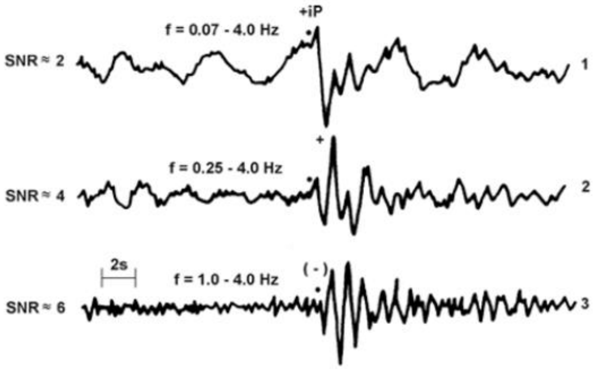
Dispersion & noise may induce phase shifts, making picking challenging (inaccurate, ambiguous association)



Black: unfiltered waveform (60 s – 50 Hz)
Red: Butterworth bandpass filter (2-10 Hz)

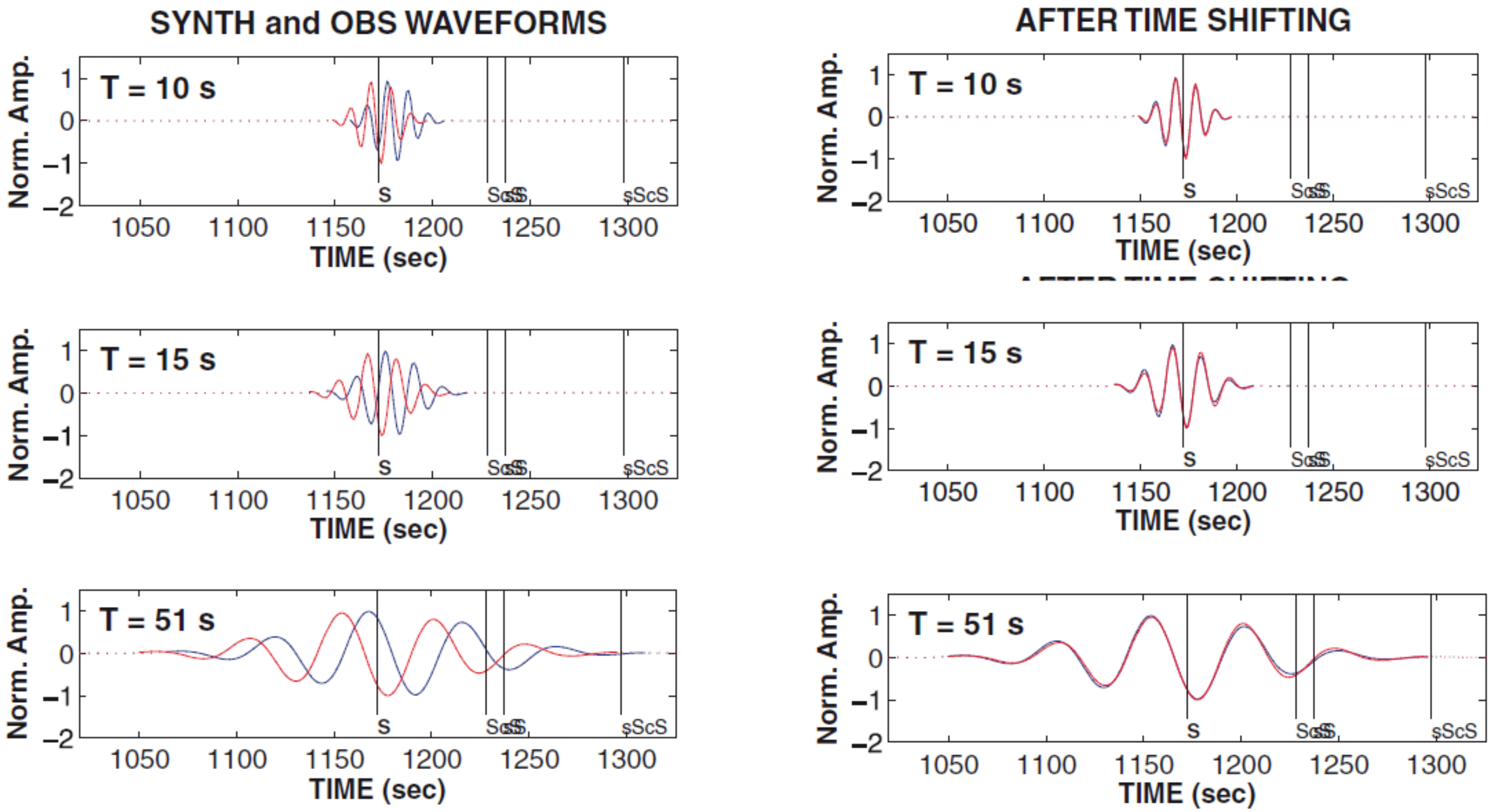
Two onset picks (green and blue) …

(From Küperkoch et al, 2010)



Wavegroup delays: cross-correlation

Confronting observed and synthetic records



Time delays by cross-correlation are frequency/period dependent (at least at long periods)

Amplitude sensitive:
well-calibrated instrument !

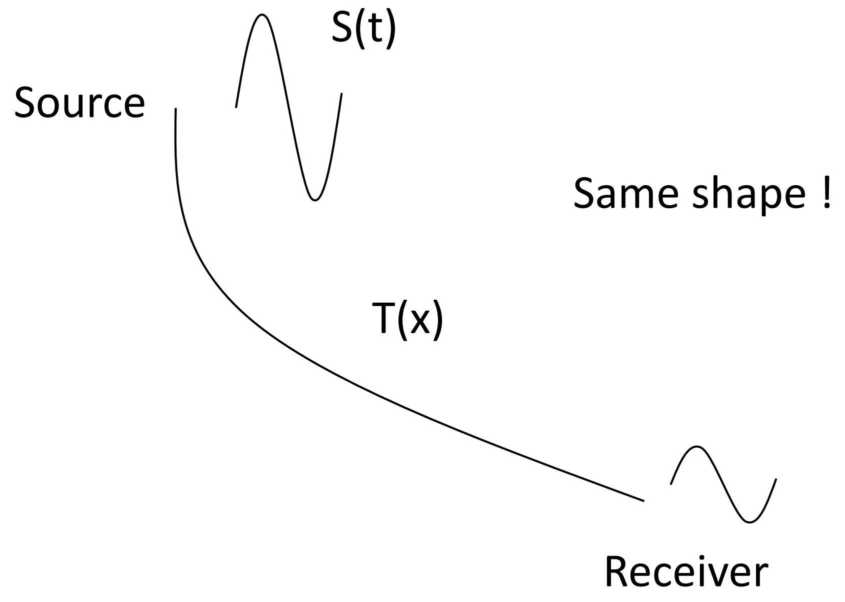
(Zaroli et al, GJInt, 2010)

Outline on first-arrival traveltime tomography

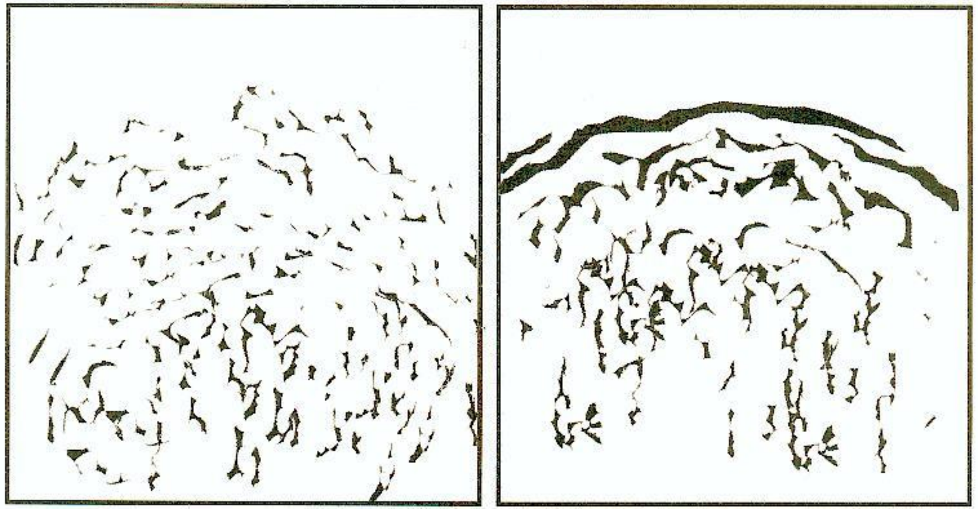
- Images at very different scales
- Waves and Phases: various concepts
- Few points on first-break ray-based tomography
- Illustration on 30-years Western Alps tomography
- First-break eikonal-based tomography
- First-break wave-equation-based tomography
- Hypocenter-velocity joint inversion
- Conclusion

Asymptotic solution: ray concept

$$u(x,t) = A(x)\, S\big(t - T(x)\big)$$
$$u(x,\omega) = A(x)\, S(\omega)\, e^{i\omega T(x)}$$



Travel-time $T(x)$ (phase $\omega T(x)$)
and Amplitude $A(x)$



Highly diffracting medium:
Loosing wavefront coherence!

Preserved wavefront:
spatial continuity

Asymptotic approach with growing frequencies

Diffraction still present!

Eikonal solution – GTD (J. Keller, 1962)

No diffraction at all!

Ray solution ⇔ Geometrical Optics ⇔∞ frequency

⇔Singularities topology (shadow zone)

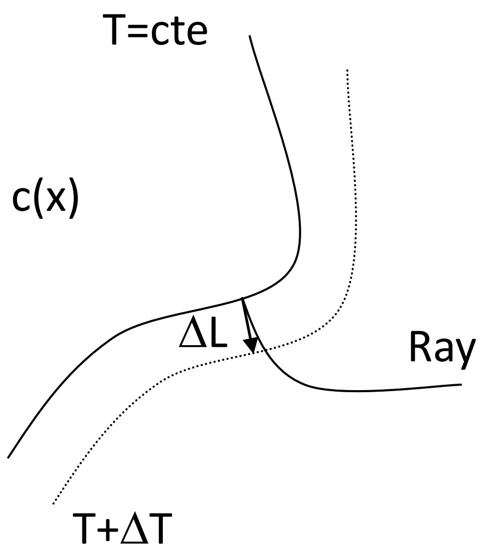
Link to the Catastrophe Theory (F. Math. René Thom)

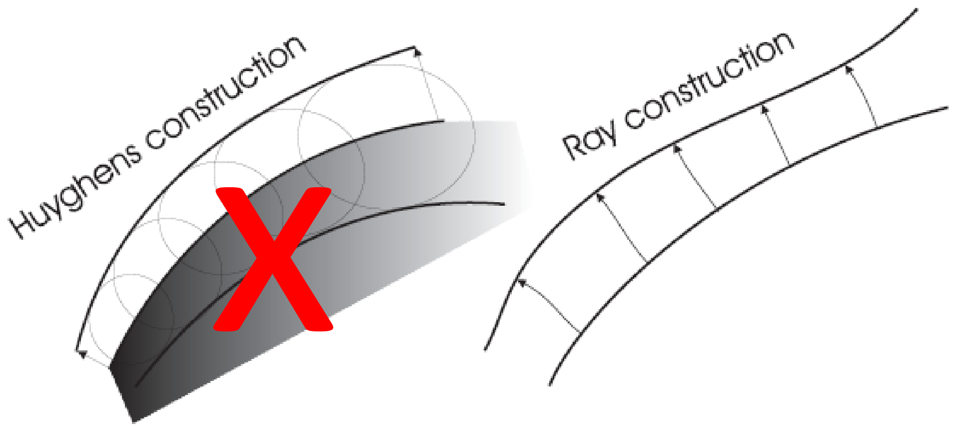
Complex phase analysis (discontinuity)…

Ray Ansatz: $A(x)e^{i\omega T(x)}S(\omega)$ - Eikonal PDE

Two simple interpretations of wavefront evolution

Orthogonal trajectories are rays in an isotropic medium





Grad(T)=$\nabla_x T$ orthogonal to wavefront

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

Direction ? : abs or square

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

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

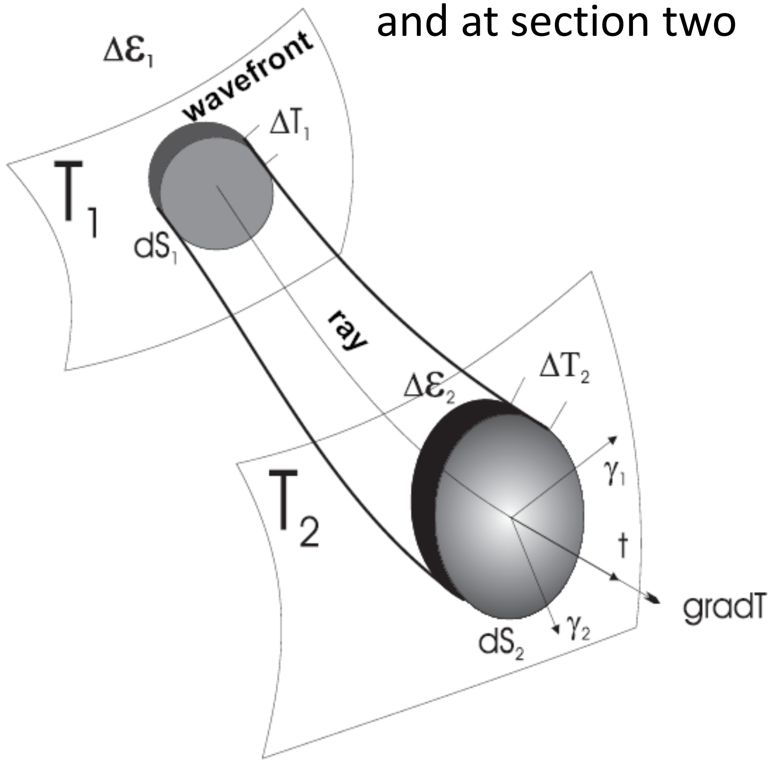
Ray Ansatz: $A(x)e^{i\omega T(x)}S(\omega)$ - Transport PDE

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

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

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

Energy flux same at section one and at section two



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

Ray Ansatz: $A(x)e^{i\omega T(x)}S(\omega)$ - Transport PDE

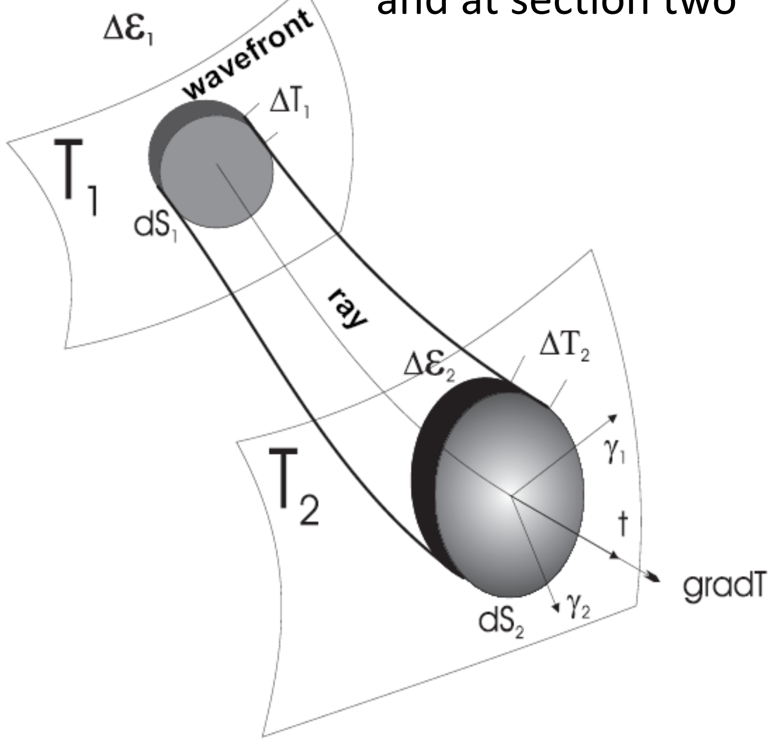
Tracing neighboring rays defines a ray tube : variation of amplitude depends on energy flux conservation through sections.

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

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

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

Energy flux same at section one and at section two



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

Ray Ansatz: $A(x)e^{i\omega T(x)}S(\omega)$ - Transport PDE

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

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

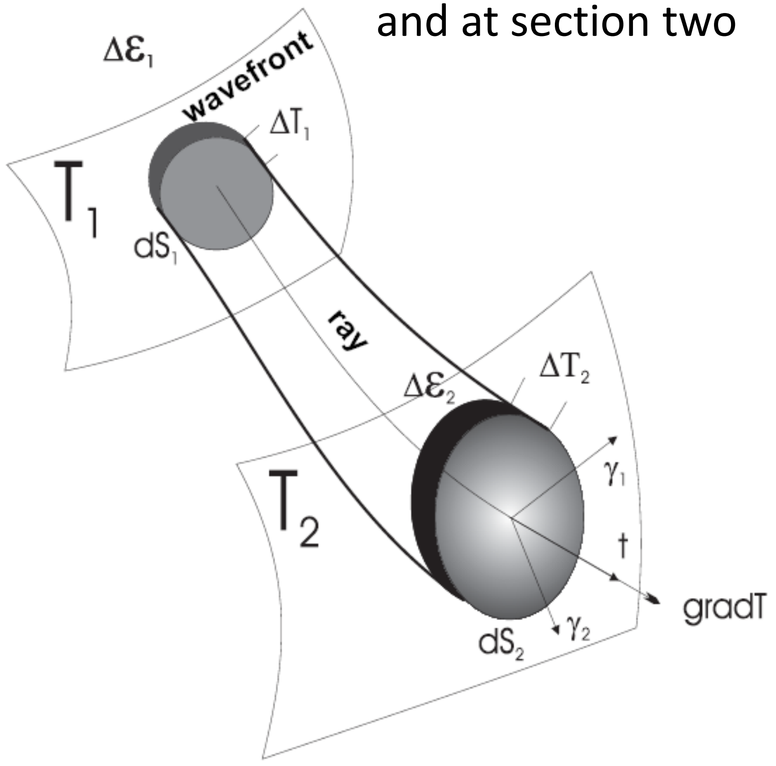
$$\rightarrow {A_1}^2 \nabla T_1 . \vec{n} dS_1 = {A_2}^2 \nabla T_2 . \vec{n} dS_2$$

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

Outward normal orientation

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

Energy flux same at section one and at section two



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

Ray Ansatz: $A(x)e^{i\omega T(x)}S(\omega)$ - Transport PDE

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

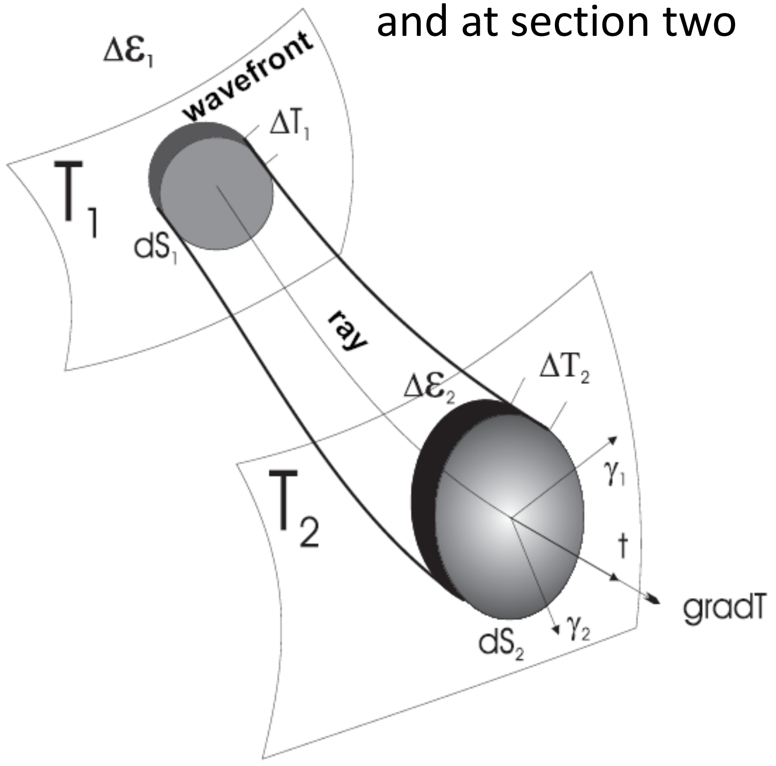
$$\Delta\mathcal{E}_1 = {A_1}^2 dS_1 \Delta T_1 = {A_2}^2 dS_2 \Delta T_2 = \Delta\mathcal{E}_2$$

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

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

Adding the flux contribution along the tube which is zero

$$0 = \iint_{Tube}^{Ray} \nabla T_c . \vec{n_c} dS_c$$



$$\boxed{2\nabla A(x) . \nabla T(x) + A(x) \nabla^2 T(x) = 0}$$

Ray Ansatz: $A(x)e^{i\omega T(x)}S(\omega)$ - Transport PDE

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

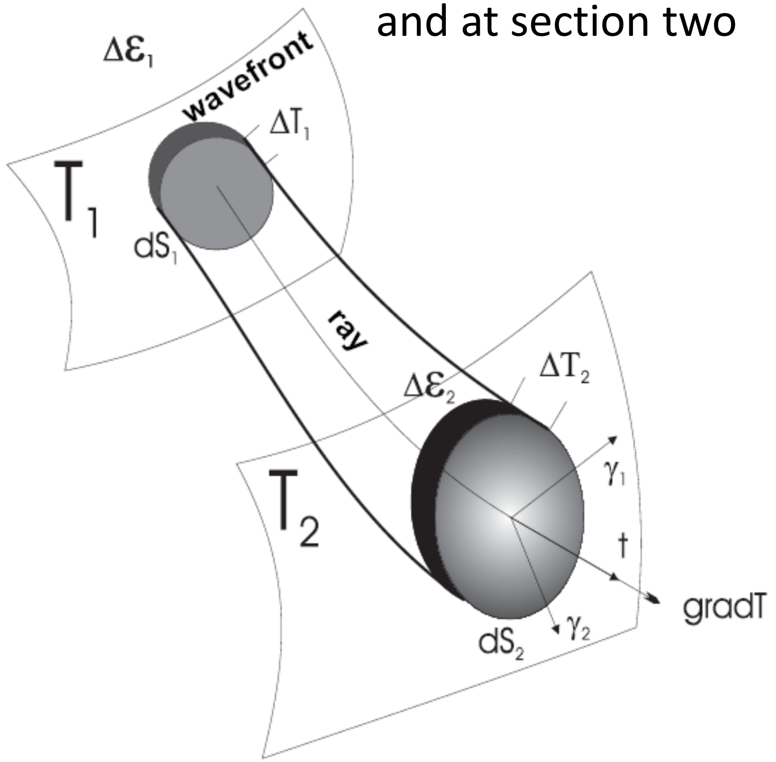
$$\Delta\mathcal{E}_1 = {A_1}^2 dS_1 \Delta T_1 = {A_2}^2 dS_2 \Delta T_2 = \Delta\mathcal{E}_2$$

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

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

$$0 = \iint_{Section}^{Ray} {A_1}^2 \nabla T_1 . \vec{n'} dS_1 + {A_2}^2 \nabla T_2 . \vec{n} dS_2 + \iint_{Tube}^{Ray} \nabla T_c . \vec{n_c} dS_c$$

Energy flux same at section one and at section two

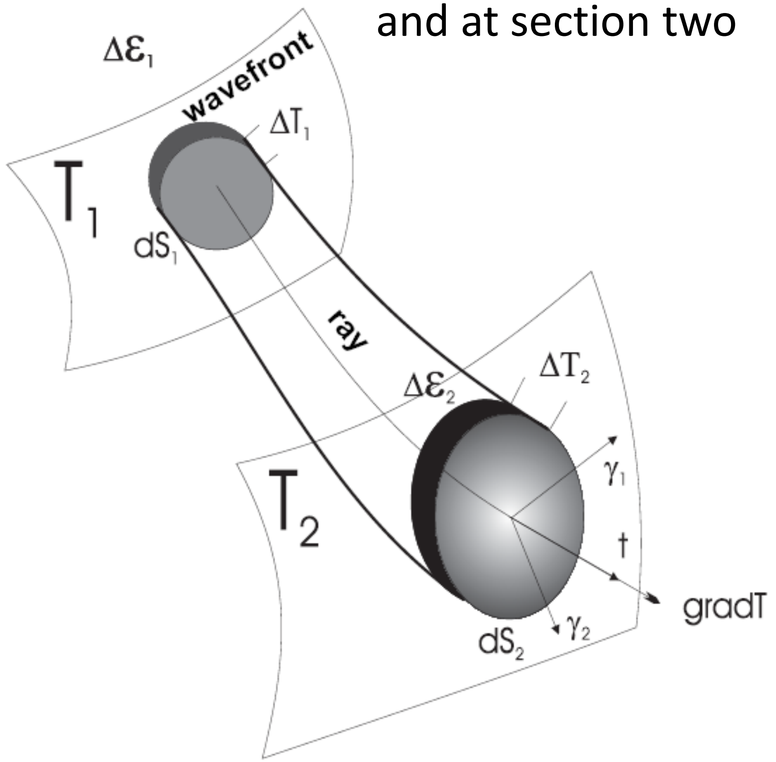


$$2\nabla A(x) . \nabla T(x) + A(x) \nabla^2 T(x) = 0$$

Ray Ansatz: $A(x)e^{i\omega T(x)}S(\omega)$ - Transport PDE

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



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

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

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

$$0 = \iint_{Section}^{Ray} {A_1}^2 \nabla T_1 . \overrightarrow{n'} dS_1 + {A_2}^2 \nabla T_2 . \vec{n} dS_2 + \iint_{Tube}^{Ray} \nabla T_c . \overrightarrow{n_c} dS_c$$

Application of the volume integral

$$0 = \iiint_{Volume}^{Ray} div(A^2\, \nabla T) dV \Longrightarrow div(A^2 \nabla T) = 0$$

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

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

A more mathematical approach

Scalar wave equation: $\nabla^2 p(x,t) = \frac{1}{c(x)} \ddot{p}(x,t)$ with velocity $c(x)$ and a constant density

Fourier transform: $p(x,t) = \exp[-i\omega t]\, p(x,\omega)$ with the circular frequency ω

Ray Ansatz: an approximate high-frequency solution: $p(x,\omega) = A(x) e^{i\omega T(x)}$
The quantities A(x) and T(x) are presumably smooth scalar functions of coordinates

$$\nabla^2 p = \nabla.\nabla p = \{i\omega(\nabla A + i\omega A \nabla T).\nabla T + (\nabla^2 A + i\omega \nabla T.\nabla A + i\omega A \nabla^2 T)\} e^{i\omega T}$$

$$-\omega^2 A\left[(\nabla T)^2 - \frac{1}{c^2}\right] + i\omega[2\nabla A.\nabla T + A\nabla^2 T] + \boldsymbol{\nabla^2 A} = 0;\ \forall\omega$$

$$(\nabla T)^2 - \frac{1}{c^2} = 0$$ Eikonal equation: non-linear PDE

$$2\nabla A.\nabla T + A\nabla^2 T = \nabla(A^2 \nabla T) = 0$$ Transport equation: linear PDE

Ray-series solution

What to do with the term $\boldsymbol{\nabla^2 A}$? Often considered as a corrective term

Amplitude may be considered as a series in inverse powers of frequency

$$A(x,\omega) = A_0(x) + \frac{1}{i\omega} A_1(x) + \frac{1}{(i\omega)^2} A_2(x)$$

Remark: fractional power
Needed for advanced concept?

No more the term $\boldsymbol{\nabla^2 A}$ but a series of transport equations for terms $A_l(x)$

Many other interpretations: hyper-eikonal or frequency-dependent eikonal

Very few applications with these transport equations

Methods of (bi)characteristics

Differential geometry (Courant & Hilbert, 1966)

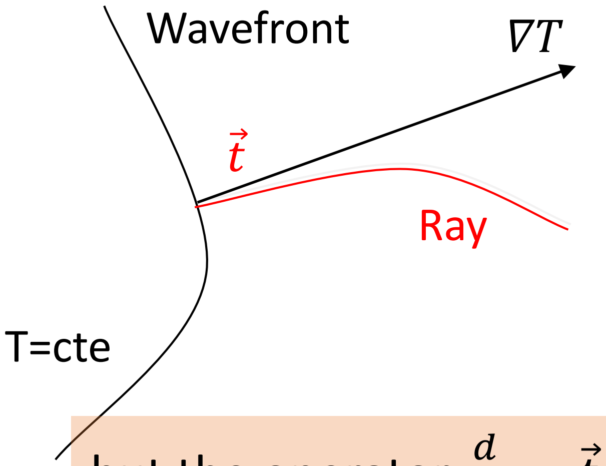
Wavefronts are geometrical characteristics

- ❖Non-linear ordinary differential equations
- ❖Lagrangian formulation as we integrate along rays (bicharacteristics)

In opposition to Eulerian formulation where we compute quantities at fixed positions

or mixed Eulerian-Lagrangian formulation using Eulerian wavefront tracking with Lagrangian ray evolution

Ray equations: bicharacteristics



$$\overrightarrow{x(s)} \; with \; s \; as \; curvilinear \; abscisse \; \vec{t} = \frac{d\vec{x}}{ds} \rightarrow \|\vec{t}\| = 1$$

- Evolution of $\vec{x}$ is given by $\frac{d\vec{x}}{ds}$: $\frac{d\vec{x}}{ds} \parallel \overrightarrow{\nabla T} \rightarrow$ $\boxed{\frac{d\vec{x}}{ds} = c(x)\overrightarrow{\nabla T}(x)}$

but the operator $\frac{d}{ds} = \vec{t}.\vec{\nabla} \;= c\overrightarrow{\nabla T}.\vec{\nabla}$

and, therefore, $\frac{d\overrightarrow{\nabla T}}{ds} = c\overrightarrow{\nabla T}.\vec{\nabla}\left(\overrightarrow{\nabla T}\right)$

leading to $\frac{d\overrightarrow{\nabla T}}{ds} = \frac{c}{2}\vec{\nabla}\left(\overrightarrow{\nabla T}.\overrightarrow{\nabla T}\right) = \frac{c}{2}\vec{\nabla}(\overrightarrow{\nabla T})^2 = \frac{c}{2}\vec{\nabla}(\frac{1}{c^2})$

Ray equations

- Evolution of $\overrightarrow{\nabla T}$ is given by $\frac{d\overrightarrow{\nabla T}}{ds}$ $\rightarrow$ $\boxed{\frac{d\overrightarrow{\nabla T}}{ds} = \vec{\nabla}(\frac{1}{c(x)})}$

Remark: characteristics are wavefronts

Ray equations: position and slowness

Two quantities along the ray:

- slowness vector $\vec{p} = \overrightarrow{\nabla T}(s)$
- position vector $\vec{q} = \vec{x}(s)$

Slowness can be removed, leading to the so-called curvature equation

$$\frac{d}{ds}\left(\frac{1}{c(s)}\frac{d\vec{q}}{ds}\right) = \vec{\nabla}\left(\frac{1}{c(s)}\right)$$

Various non-linear ray equations: which ODE to choose for integration? Time or Particule

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

$$\frac{d\vec{q}(t)}{dt} = c^2(\vec{q})\vec{p}$$
$$\frac{d\vec{p}(t)}{dt} = c(\vec{q})\nabla\frac{1}{c(\vec{q})}$$

Particule stepping

$$\frac{d\vec{q}(\xi)}{d\xi} = \vec{p}$$
$$\frac{d\vec{p}(\xi)}{d\xi} = \frac{1}{c(\vec{q})}\nabla\frac{1}{c(\vec{q})}$$
$$\frac{dT(\xi)}{d\xi} = \frac{1}{c^2(\vec{q})}$$

The simplest set

Ray equations: position and slowness

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

$$\frac{d\vec{q}(t)}{dt} = c^2(\vec{q})\vec{p}$$
$$\frac{d\vec{p}(t)}{dt} = c(\vec{q})\nabla \frac{1}{c(\vec{q})}$$

$$dt = \frac{1}{c(\vec{q})} ds = \frac{1}{c(\vec{q})^2} d\xi$$

Particule stepping

$$\frac{d\vec{q}(\xi)}{d\xi} = \vec{p}$$
$$\frac{d\vec{p}(\xi)}{d\xi} = \frac{1}{c(\vec{q})} \nabla \frac{1}{c(\vec{q})}$$
$$\frac{dT(\xi)}{d\xi} = \frac{1}{c^2(\vec{q})}$$

Any numerical integration tool: Runge-Kutta or Predictor-Corrector schemes. However, Eikonal quantity $p^2 = 1/c^2(\vec{q})$ may be used for quality control. No need of automatic control of schemes.

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

Anisotropic media: ray equations

Particule stepping

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

Eikonal equation

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

Numerical tools can be used as well, including the perturbation theory for paraxial ray equations

Time stepping

$$\frac{dq_i\,(t)}{dt} = (p_i + \frac{1}{c^3(q_i,p_i)}\frac{\partial c(q_i,p_i)}{\partial p_i})\frac{1}{\sum_k p_k(p_k + \frac{1}{c^3(q_k,p_k)}\frac{\partial c(q_k,p_k)}{\partial p_k})}$$
$$\frac{dp_i\,(t)}{dt} = (\frac{1}{c(q_i,p_i)}\nabla_{q_i}\frac{1}{c(q_i,p_i)})\frac{1}{\sum_k p_i(p_k + \frac{1}{c^3(q_k,p_k)}\frac{\partial c(q_k,p_k)}{\partial p_k})}$$

Anisotropic media: transport equations

Isotropic case: scalar functions

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

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

$$u(x,t) = A(x)\, S\big(t - T(x)\big)$$
$$\mathrm{u}(\mathrm{x},\omega) = \mathrm{A}(\mathrm{x})\, \mathrm{S}(\omega)\, e^{i\omega T(x)}$$

Anisotropic case: vectorial functions

$$\vec{\mathrm{u}}(\mathrm{x},\mathrm{t}) = \vec{\mathrm{A}}(\mathrm{x})\, \mathrm{S}\big(\mathrm{t} - \mathrm{T}(\mathrm{x})\big)$$
$$\vec{\mathrm{u}}(\mathrm{x},\omega) = \vec{\mathrm{A}}(\mathrm{x})\, \mathrm{S}(\omega)\, \mathrm{e}^{\mathrm{i}\omega\mathrm{T}(\mathrm{x})}$$

$\forall i = 1,3$

$$\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}\left(\frac{\partial}{\partial x_j}\big(c_{ijkl}(x)\big)\frac{\partial A_k}{\partial x_k} + c_{ijkl}(x)\frac{\partial^2 A_k}{\partial x_j \partial x_l}\right) = 0$$

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

$$\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}(x)A_k\frac{\partial T}{\partial x_j}\frac{\partial T}{\partial x_l} - \rho(x)A_i = 0$$

Slawinski, 2003, p166-167

Not easy to understand geometrically how energy flows

Solving ray equations

How to solve these equations?

Ordinary Differential Equations

Lagrangian formulation (FAST!)

Wavefront complexity

Simple properties of these ODEs

- Intrinsic solutions independent of the coordinate sytem used to solve it
- If a dummy variable in velocity description, use it as the variable stepping (often x coordinate)

$$\nabla_x \frac{1}{c(q_z)} = 0 \Rightarrow p_x = cte \Rightarrow q_x = {q_x}^0 + \xi p_x \text{ (1)}$$

- 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

Velocity variation $v(z)$

Ray equations are

$$\frac{dq_x}{d\tau} = p_x; \frac{dq_y}{d\tau} = p_y; \frac{dq_z}{d\tau} = p_z$$

$$\frac{dp_x}{d\tau} = 0; \frac{dp_y}{d\tau} = 0; \frac{dp_z}{d\tau} = u(z)\frac{du(z)}{dz}$$

$$u(z) = \frac{1}{v(z)}$$

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.

$$\frac{dq_x}{dq_z} = \frac{p_x}{p_z} = \frac{p_x}{\pm\sqrt{u^2(z) - p_x^{\ 2}}}$$

where p_x is a constante

$$\frac{dq_x}{d\tau} = p_x; \frac{dq_z}{d\tau} = p_z$$

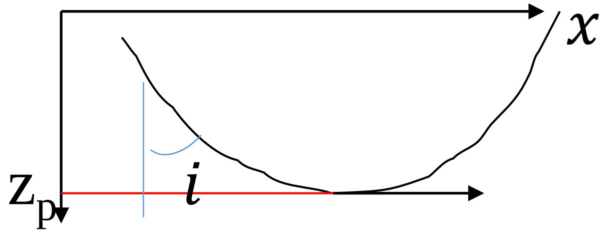
$$\frac{dp_x}{d\tau} = 0; \frac{dp_z}{d\tau} = u(z)\frac{du(z)}{dz}$$

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

quadratic expression wrt depth

Velocity variation $v(z)$

At a given maximum depth z_p, the slowness vector is horizontal following the equation $p_x^2 = p^2 = u^2(z_p)$



$$q_x(z_1,p_1) = q_x(z_0,p_0) + \int_{z_0}^{z_p} \frac{p_x}{\sqrt{u^2(z) - p_x^2}} dz + \int_{z_1}^{z_p} \frac{p_x}{\sqrt{u^2(z) - p_x^2}} dz$$

$$T(z_1,p_1) = T(z_0,p_0) + \int_{z_0}^{z_p} \frac{u^2(z)}{\sqrt{u^2(z) - p_x^2}} dz + \int_{z_1}^{z_p} \frac{u^2(z)}{\sqrt{u^2(z) - p_x^2}} dz$$

Simple quadratic equations

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

In Cartesian frame

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

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

$$with\ p = u\ sini$$

In Spherical frame

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

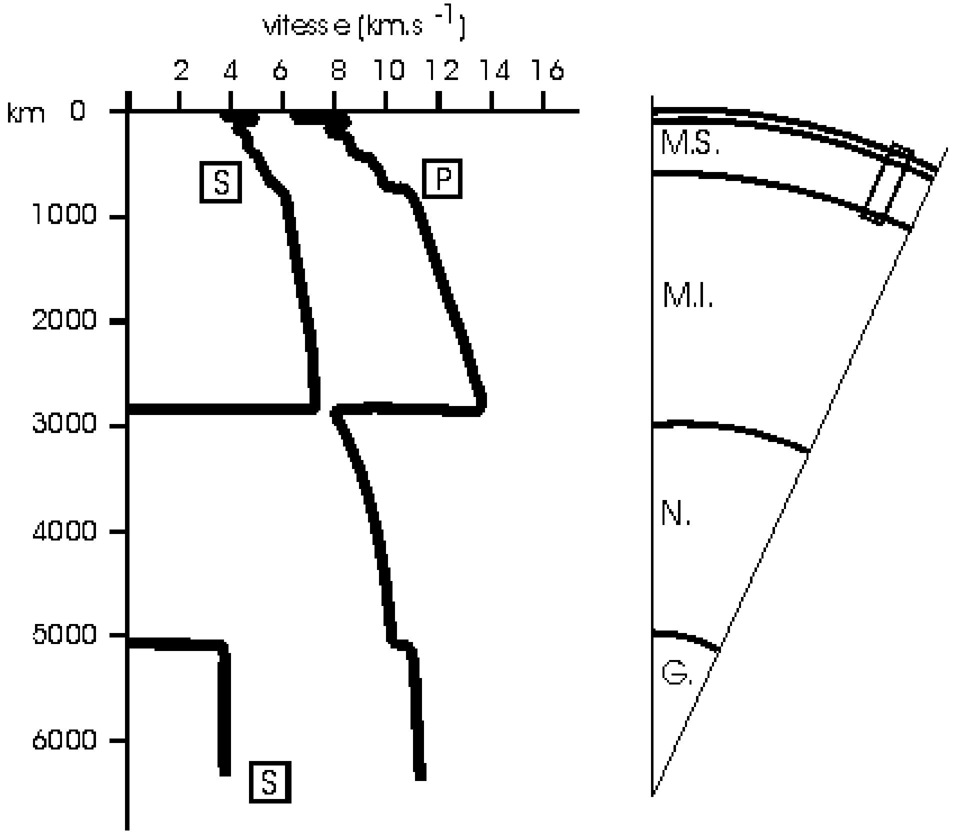
$$T(p) = 2 \int_{r_p}^{a} \frac{p}{\sqrt{r^2u^2(r) - p^2}} \frac{dr}{r}$$

$$with\ p = ru\ sini$$

Velocity structure in the Earth

Radial Structure

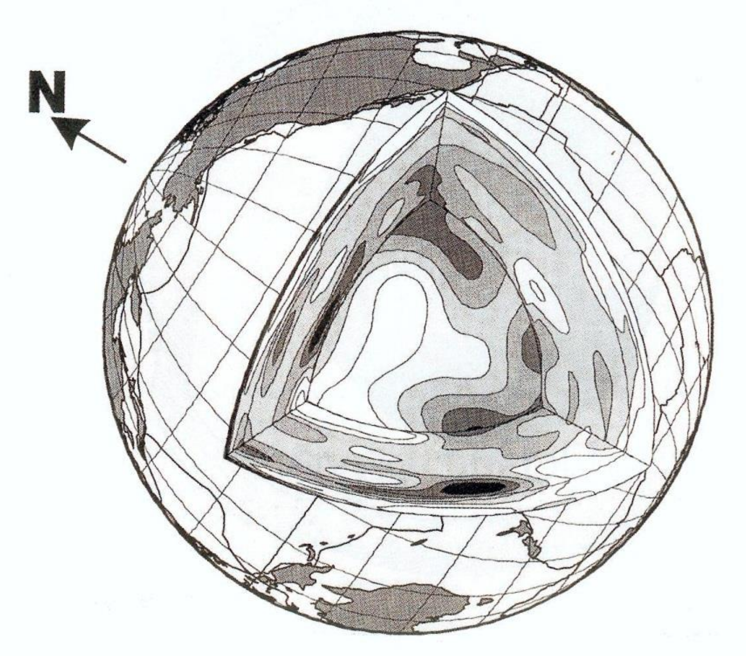
PREM – AKI135 models…



(a)

What about velocity $v(x, y, z)$

No quadratic formulation!



Numerical approaches
for solving differential equations.

Solving ray equations when v(x,y,z)

How to solve these equations?

Ordinary Differential Equations

Lagrangian formulation (still FAST!)

Wavefront complexity

Hamilton's ray equations

$$\frac{d\vec{q}(\xi)}{d\xi} = \vec{p}$$
$$\frac{d\vec{p}(\xi)}{d\xi} = \frac{1}{c(\vec{q})}\nabla_{\vec{q}}\frac{1}{c(\vec{q})}$$

$$\mathcal{H}(\vec{q},\vec{p}) = \frac{1}{2}\left(p^2 - \frac{1}{c^2(\vec{q})}\right) = 0$$

$$\frac{d\vec{q}(\xi)}{d\xi} = \nabla_{\vec{p}}\mathcal{H}$$
$$\frac{d\vec{p}(\xi)}{d\xi} = -\nabla_{\vec{q}}\mathcal{H}$$
$$\frac{dT}{d\xi} = \vec{p}.\nabla_{\vec{p}}\mathcal{H}$$

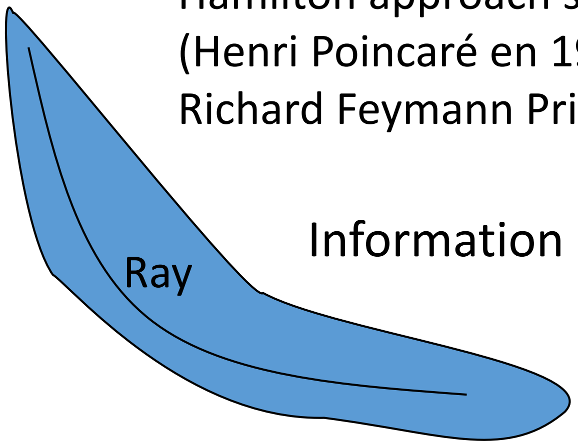
Mechanics : ray tracing as a particular balistic problem

sympletic structure (FUN!)

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

$$\vec{q}_0 + \delta\vec{q}$$
$$\vec{p}_0 + \delta\vec{p}$$

$$\delta q \text{ and } \delta p \text{ "small"}$$



Information around the ray

Meaning of the neighborhood zone
Fresnel zone if finite frequency
Any zone depending on your problem Gaussian Beam summation

Reduced Hamilton formulation

Reduction of ray equations from six to four (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

Hamilton’s ray equations

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)

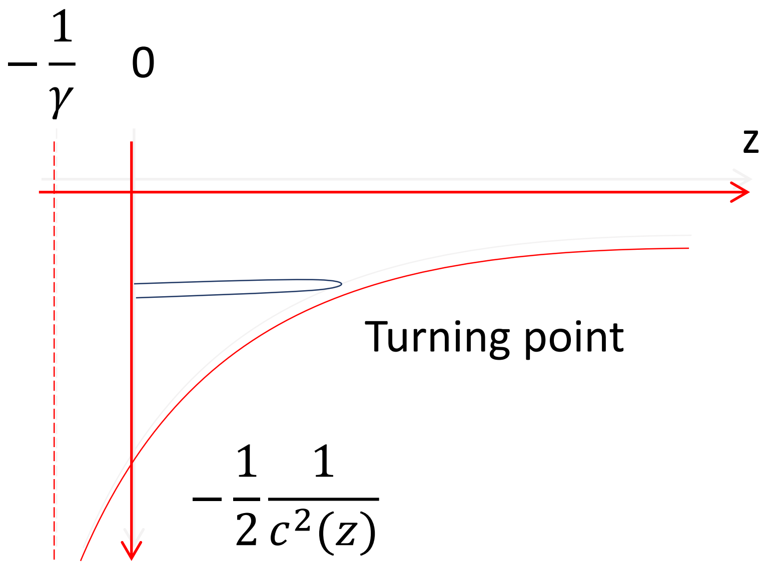
- Computational complexity for number crunching
- No evaluation in the vicinity of the hyper-surface of the eikonal conservation as we live in lower-dimension space

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



Kinetic energy

Potential energy

$$\mathcal{H}(\vec{q},\vec{p}) = \frac{1}{2}\frac{p^2}{m} + V(\vec{q})$$

$$\mathcal{H}(\vec{q},\vec{p}) = \frac{1}{2}p^2 - \frac{1}{2}\frac{1}{c^2(\vec{q})}$$

$$c(z) = c_0(1+\gamma z)$$

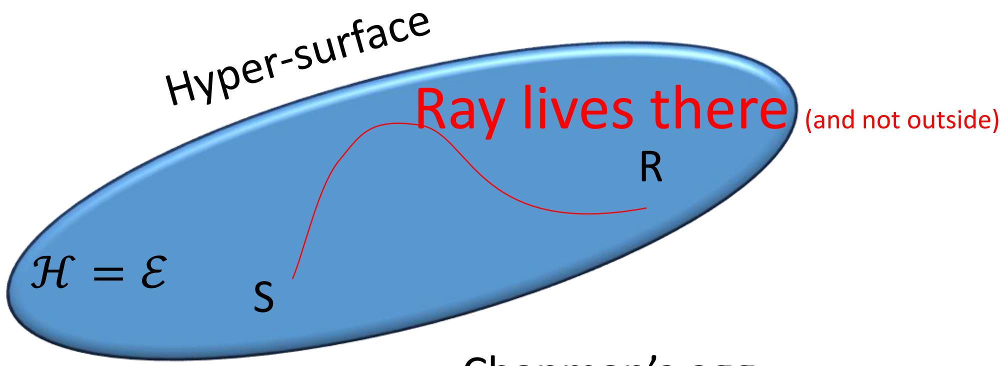
ray is a circle

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

A conservative system

$$\mathcal{H}(\vec{q},\vec{p}) = 0$$

Just FUN?



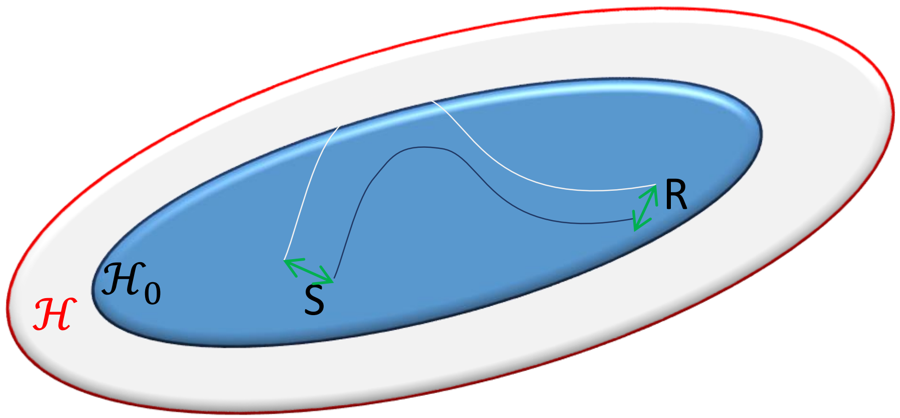
Chapman's egg

A non-conservative system

$$\mathcal{H}(\vec{q},\vec{p}) = \mathcal{E}(\vec{q},\vec{p}) \Rightarrow \mathcal{H}(\vec{q},\vec{p}) - \mathcal{E}(\vec{q},\vec{p}) = 0$$

Embedding it into an isolated system !

Mechanical point of view



Just FUN: NO!

If perturbation of the velocity structure, should we reset ray tracing?

$$\mathcal{H}(\vec{q},\vec{p}) = \mathcal{H}_0(\vec{q},\vec{p}) + \Delta\mathcal{H}(\vec{q},\vec{p})$$

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.

Low complexity of ray equations

Keeping computer complexity low!

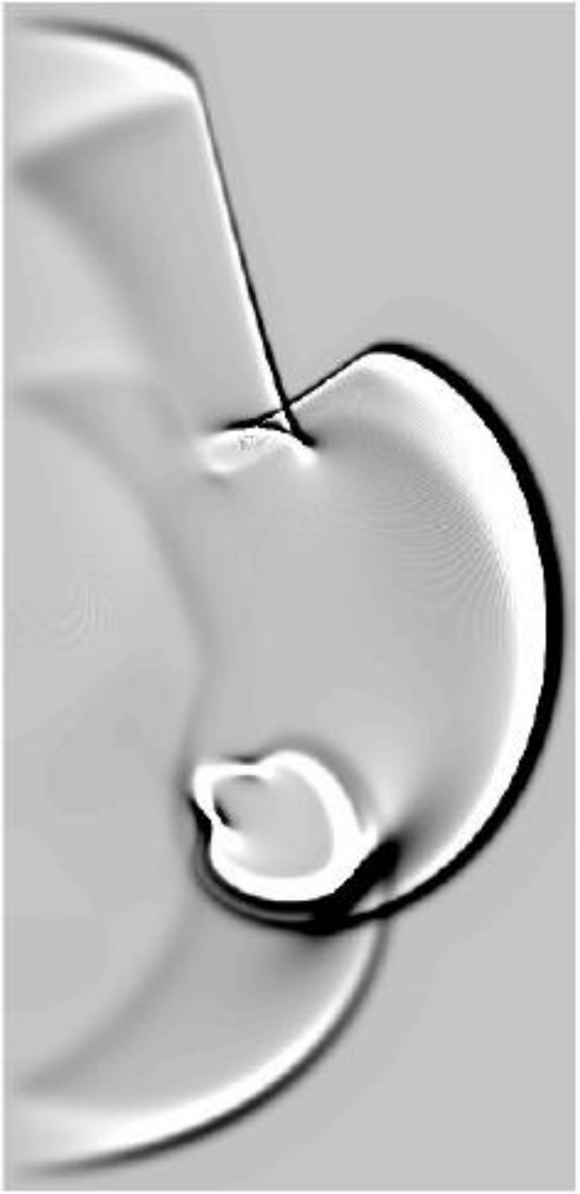
- Ray tracing is a fast 1D integration in 2D/3D

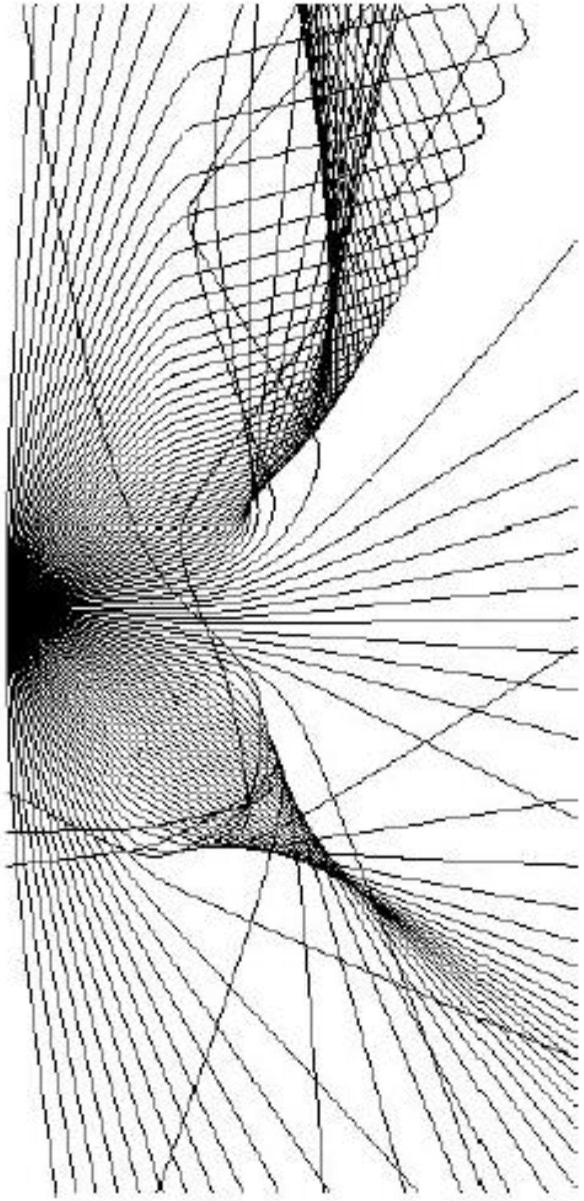
- Ray tracing equations as ODEs may sample the model quite evenly

- Lagrangian formulation: we follow a point while tracing rays without regarding the density of rays inside the model

Interpolation and extrapolation challenges

Wave solution





Ray solution

Interpolation and extrapolation challenges

How to control the ray sampling of the model?

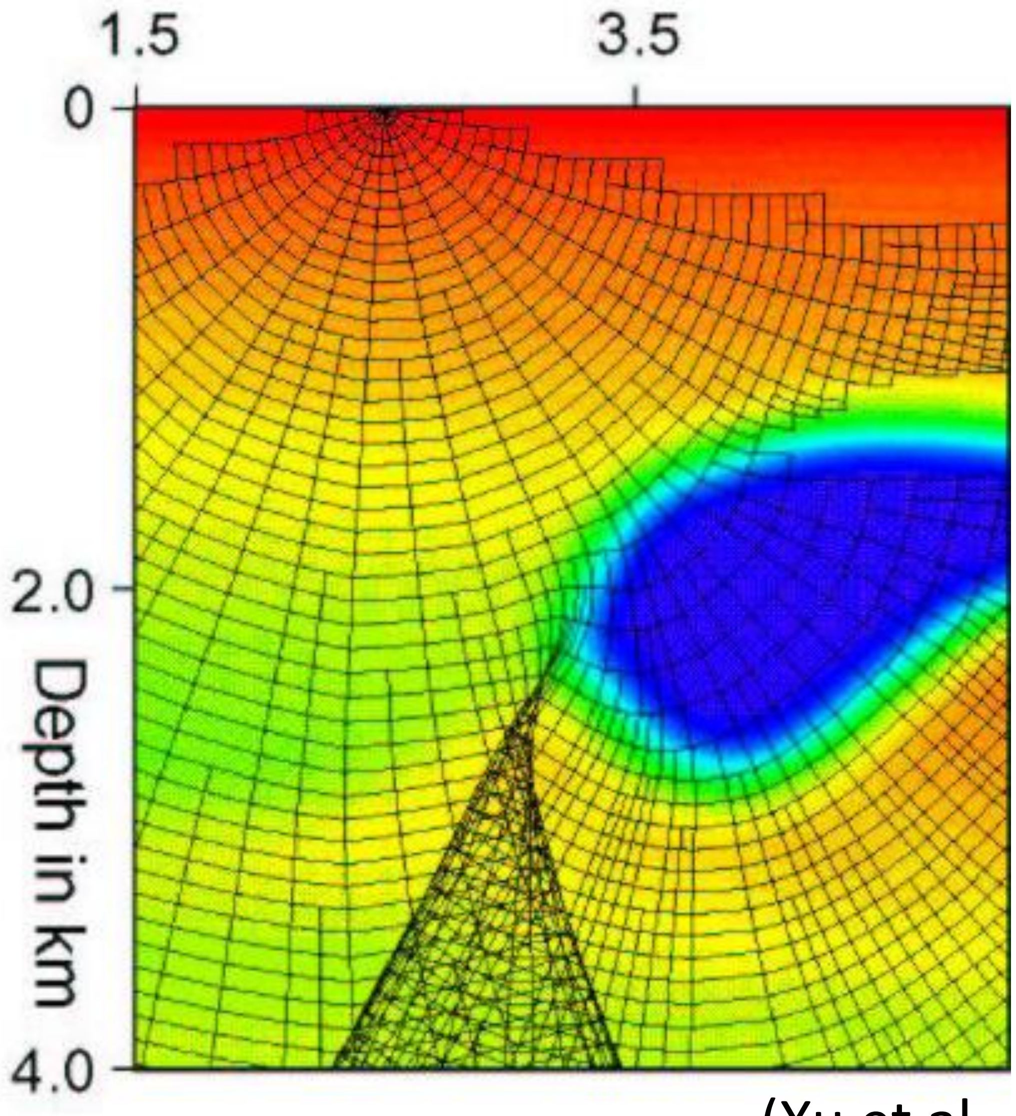
- ❑ Folding zones!

Available information

- ❑ Shadow zones!

Missing information

Semi-lagrangian approach



(Xu et al., 1999)

Tracking the wavefront and its complexity with rays: still shadow zones … but allows the folding of rays.

Adding and suppressing rays for correct sampling of the wavefront!

HOW ?

Providing separated elemental contributions

(Jin et al, 1992; Lambaré et al, 1992)

(Vinje et al, 1993)

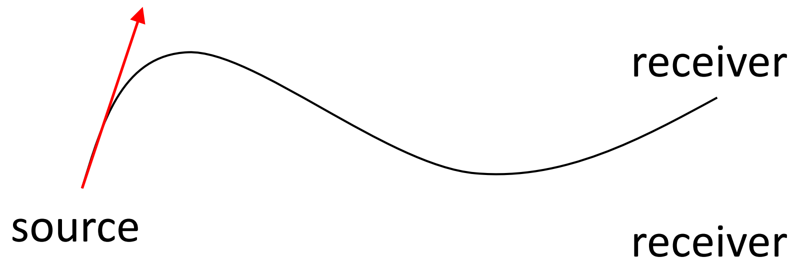
Integration of ray equations

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

Runge-Kutta second-order integration

Predictor-Corrector integration

A very good QC: the eikonal must be equal to zero !



Initial conditions EASY

Boundary conditions VERY DIFFICULT

Shooting δp ?

Bending δx ?

Continuing δc ?

Save slowness conditions if possible !

AND FROM TIME TO TIME IT FAILS ! (inherent to geometrical optics)

But we need 2-points ray tracing because we have a source and a receiver to connect ! We even need more: branch identification (triplication for example)

Non-linear ray tracing

Runge-Kutta integration

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

$$f^{1/2} = f^0 + \frac{\Delta\xi}{2} A(f^0)$$

Second-order RK integration $$f^1 = f^0 + \Delta\xi \ A\left(f^{1/2}\right)$$

Fourth-order RK integration (home work!)

Predictor-Corrector integration?

Show simple toy python codes for doing so.

ODE building

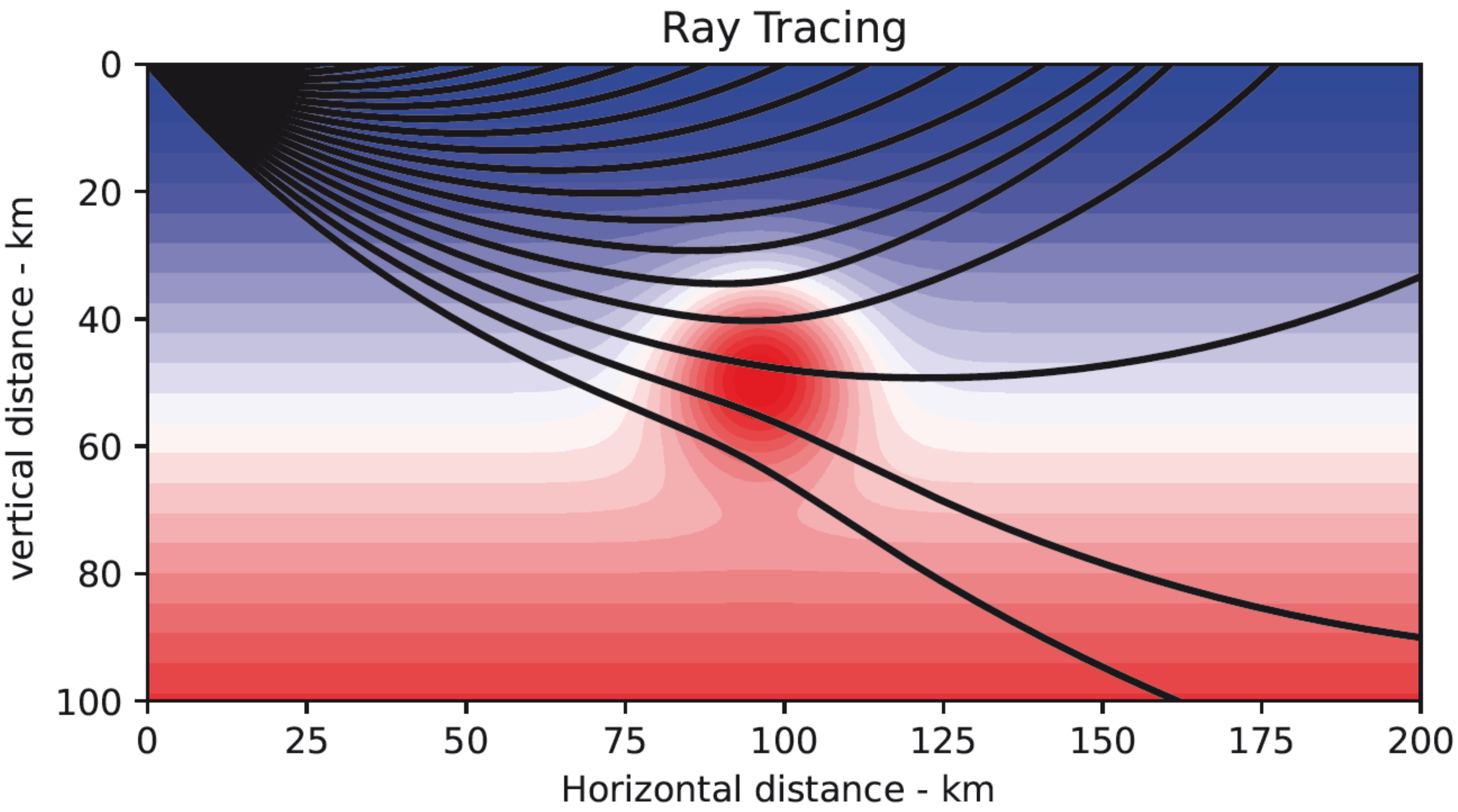
- Runge-Kutta of second order
- Write a computer program for an analytical law for the velocity: take a gradient with a component along x and a component along z

Home work : redo the same thing with a Runge-Kutta of fourth order (look after its definition) and predictor-corrector scheme.

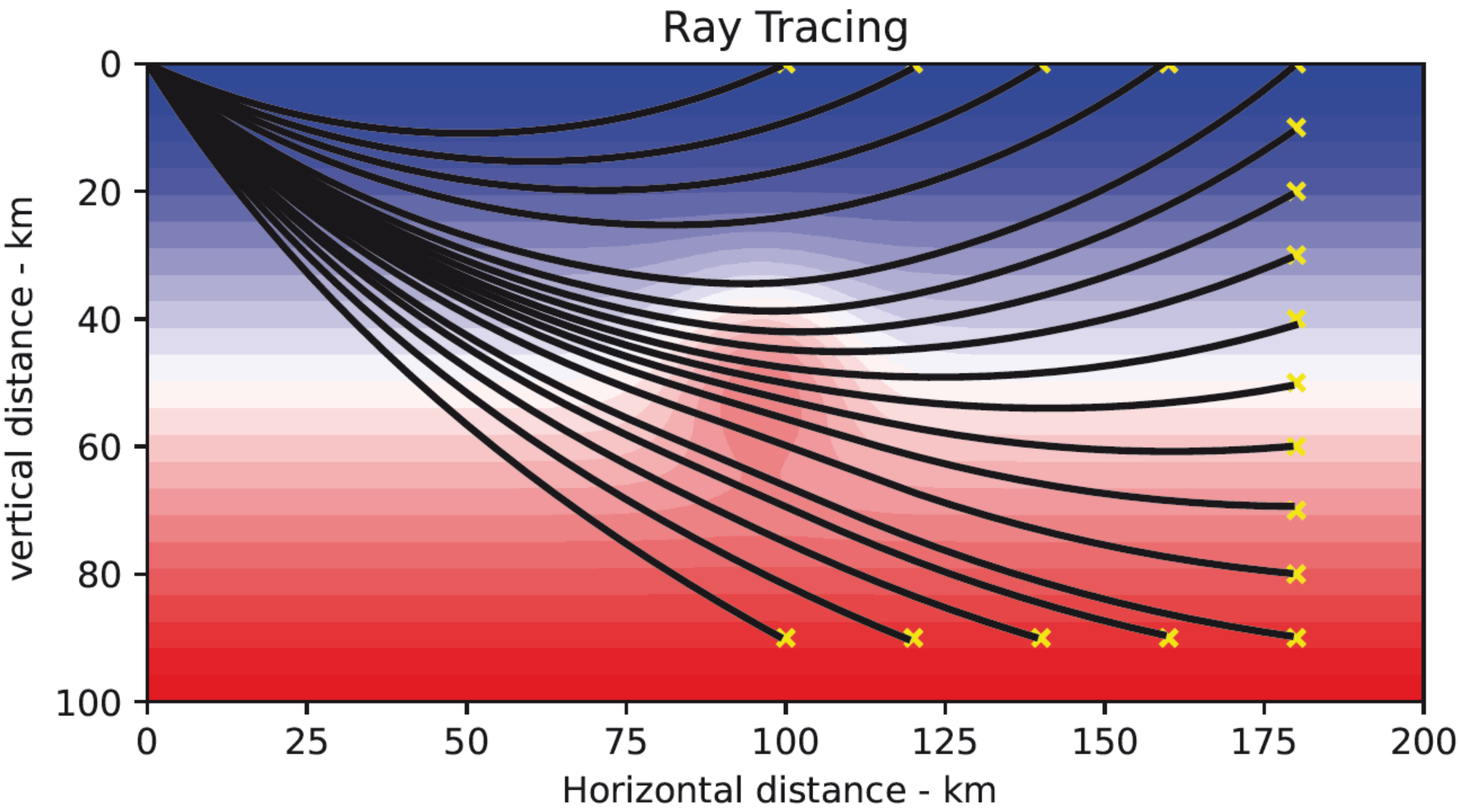
Consider a gradient of the square of slowness and/or a vertical gradient of velocity

Consider a model defined by a grid with spline interpolation for computing spatial derivatives

Ray tracing with initial conditions



Two-points ray tracing: how ?



Boundary conditions

How to sample the model
around a given ray?

How to consider interfaces?

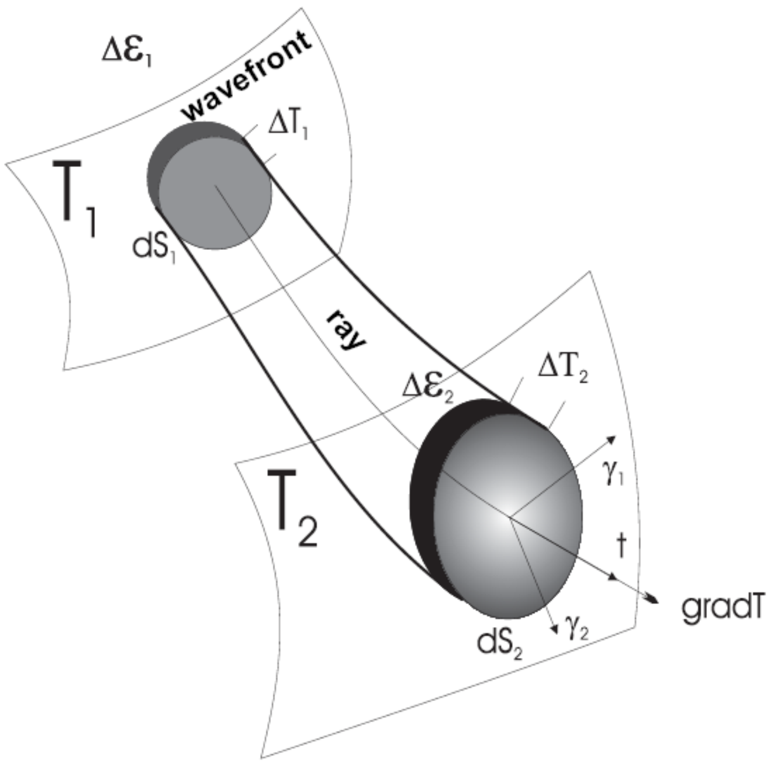
Paraxial theory: what happens around one ray?

Paraxial ray theory

similar to

Gauss optics

Paraxial equations: amplitude estimation



Ray tube allows amplitude estimation

No need to solve the transport equation!

One has to compute rays nearby the current ray for an estimation of the ray tube.

Keeping a ray in the neighbooring of the so-called central ray is quite difficult when using ray equations with Lagrangian formulation.

Perturbation theory is the way to go and tools from mechanics can be used …

and many other things…

Paraxial equations: amplitude estimation

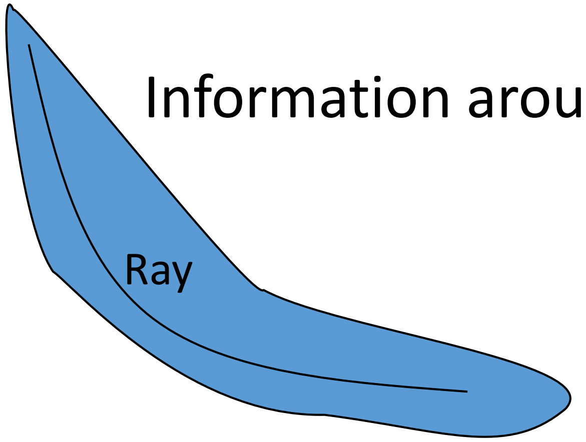
Hamilton's ray equations

$$\frac{d\vec{q}(\xi)}{d\xi} = \vec{p}$$
$$\frac{d\vec{p}(\xi)}{d\xi} = \frac{1}{c(\vec{q})}\nabla_{\vec{q}}\frac{1}{c(\vec{q})}$$

$$\mathcal{H}(\vec{q},\vec{p}) = \frac{1}{2}\left(p^2 - \frac{1}{c^2(\vec{q})}\right)$$

$$\frac{d\vec{q}(\xi)}{d\xi} = \nabla_{\vec{p}}\mathcal{H}$$
$$\frac{d\vec{p}(\xi)}{d\xi} = -\nabla_{\vec{q}}\mathcal{H}$$
$$\frac{dT}{d\xi} = \vec{p}.\nabla_{\vec{p}}\mathcal{H}$$

$$y_0 = \begin{pmatrix}\vec{q}\\ \vec{p}\end{pmatrix}$$ Central solution



Information around the ray y_0

$$\begin{matrix}\vec{q}_0 + \delta\vec{q}\\ \vec{p}_0 + \delta\vec{p}\end{matrix} \qquad \delta\vec{q} \text{ and } \delta\vec{p} \text{ "small"}$$

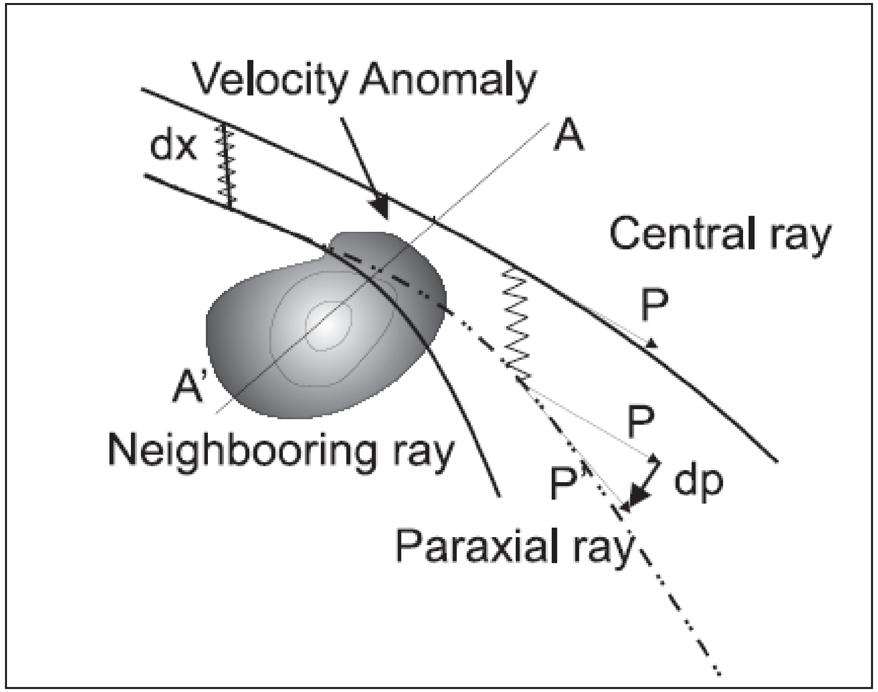
$$\delta y = \begin{pmatrix}\overrightarrow{\delta q}\\ \overrightarrow{\delta p}\end{pmatrix}$$ Perturbation solution

Paraxial ray equations

The perturbation machinery

$$\frac{d(\overrightarrow{q_0}+\overrightarrow{\delta q})}{d\xi} = \nabla_{\overrightarrow{p_0}+\overrightarrow{\delta p}}\mathcal{H}\left(\overrightarrow{q_0}+\overrightarrow{\delta q},\overrightarrow{p_0}+\overrightarrow{\delta p}\right)$$

$$\frac{d\overrightarrow{\delta q}}{d\xi} = \nabla_{\overrightarrow{p_0}}\nabla_{\overrightarrow{p_0}}\mathcal{H}(\overrightarrow{q_0},\overrightarrow{p_0})\overrightarrow{\delta p} + \nabla_{\overrightarrow{p_0}}\nabla_{\overrightarrow{q_0}}\mathcal{H}(\overrightarrow{q_0},\overrightarrow{p_0})\overrightarrow{\delta q}$$



$$\frac{d}{d\xi}\begin{bmatrix}\overrightarrow{\delta q}\\ \overrightarrow{\delta p}\end{bmatrix} = \begin{bmatrix}\nabla_{pq}\mathcal{H}^0 & \nabla_{pp}\mathcal{H}^0\\ -\nabla_{qq}\mathcal{H}^0 & -\nabla_{qp}\mathcal{H}^0\end{bmatrix}\begin{bmatrix}\overrightarrow{\delta q}\\ \overrightarrow{\delta p}\end{bmatrix}$$

$$\frac{d}{d\xi}\left(\overrightarrow{\delta y}\right) = A(\vec{y}_0)\overrightarrow{\delta y}$$

The matrix A does not depend on unknown quantities δy but only on quantities y_0: LINEAR PROBLEM (SIMPLE) !

Paraxial solution can be used for different purposes: *amplitude estimation, ray tube crossing (KMAH index), two-points ray tracing problem* and so on

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)

One can think that paraxial ray is somehow a ray derivative of the central ray. It is arbitrarily near the central ray.

2D simple linear system: isotropic case

$$\frac{d}{d\xi}\begin{bmatrix}\delta q_x\\ \delta q_z\\ \delta p_x\\ \delta p_z\end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0.5\,\frac{\partial^2\, 1/c^2(\vec{q}_0)}{\partial x^2} & 0.5\,\frac{\partial^2\, 1/c^2(\vec{q}_0)}{\partial x \partial z} & 0 & 0\\ 0.5\,\frac{\partial^2\, 1/c^2(\vec{q}_0)}{\partial z \partial x} & 0.5\,\frac{\partial^2\, 1/c^2(\vec{q}_0)}{\partial z^2} & 0 & 0 \end{bmatrix}\begin{bmatrix}\delta q_x\\ \delta q_z\\ \delta p_x\\ \delta p_z\end{bmatrix}$$

Linear system

More complex for anisotropic structure but still workable

Four elementary paraxial trajectories

NOT A paraxial RAY !

$\delta y1^t(0)=(1,0,0,0)$

$\delta y2^t(0)=(0,1,0,0)$

$\delta y3^t(0)=(0,0,1,0)$

$\delta y4^t(0)=(0,0,0,1)$

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

Any paraxial ray is a linear combination of these four elementary trajectories

Numerical integration

Second-order RK integration $$\frac{df}{d\xi} = A(f)$$

Non-linear ray tracing
$$f^{1/2} = f^0 + \frac{\Delta\xi}{2} A(f^0)$$
$$f^1 = f^0 + \Delta\xi \; A\left(f^{1/2}\right)$$

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

Linear paraxial ray tracing

Propagator technique

Optical Lens technique

$$\frac{d\delta f}{d\xi} = A(f)\delta f$$
$$\delta f^1 = \delta f^0 + \Delta\xi \; A(f^0)\delta f^0$$

2D paraxial conditions

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!

$$\delta\mathcal{H}(\xi) = \delta\mathcal{H}(0) = 0$$

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

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 condition

Point source: no shift in the position when doing perturbation:

$$\delta q_x(O) = \delta q_z(O) = 0 \;\Rightarrow\; p_x(0)\delta p_x(0) + p_z(0)\delta p_z(0) = 0$$

$$\delta p_x(0) = \alpha p_z(0)$$

$$\delta p_z(0) = -\alpha p_x(0)$$

This is enough to verify this condition initially

α *arbitrary constant (linear system)*

Point source paraxial solution $\delta y^{\alpha}(\xi) = \alpha p_z(0)\ \delta y3(\xi) - \alpha p_x(0)\ \delta y4(\xi)$ elementary trajectories

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 α could be set to an arbitrary small value: this is a derivative or plan tangent computation (Gauss optics)

Plane source condition

Plane source: shift in the shooting direction when doing perturbation:

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

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

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

α *arbitrary constant (linear system)*

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

Cerveny's condition (both x and z variation)

Paraxial solution $\delta y^b(\xi) = \alpha\,\frac{1}{2}\frac{\partial 1/c^2(x,z)}{\partial z}(0)\,\delta y1(\xi) - \alpha\,\frac{1}{2}\frac{\partial 1/c^2(x,z)}{\partial x}(0)\,\delta y2(\xi)$

We combine the first two paraxial ray trajectories.

elementary trajectories

Two independent paraxial rays in 2D (δy^a *and* δy^b): point (seismograms) and plane (beams) paraxial rays

Chapman source condition (free surface)

Chapman condition: keep the shift position along the free surface:

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

$$\delta q_x(0) = \alpha p_x(0)$$

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

Only derivative along the x direction: no extrapolation

Chapman's condition (only z variation)

$\alpha\ arbitrary\ constant\ (linear\ system)$

Paraxial solution $\delta y'(\xi) = \alpha p_x(0)\,\delta y1(\xi) - \alpha\frac{1}{2}\frac{\partial 1/c^2(x,z)}{\partial x}(0)\delta y3(\xi)$

We combine the first two paraxial ray trajectories.

elementary trajectories

Only perturbation along x which is important when considering source at the free surface

Paraxial source conditions

Two independent paraxial rays in 2D (δy^a and δy^b):
point (seismograms) and plane (beams) paraxial rays

Four independent paraxial rays in 3D ($\delta y^a, \delta y^b, \delta y^c, and\ \delta y^d$):
2 point (seismograms) and 2 plane (beams) paraxial rays

Remark: working with trajectories implies that paraxial conditions could be defined on the fly for having local conditions at different points of the model

Remark: 3 point and 3 plane paraxial trajectories in 3D!

KMAH index tracking through paraxial

KMAH index key element for seismograms

$$u(\vec{q},t) = A(\vec{q})e^{i\omega T(x)}e^{-i\frac{\pi}{2}sgn(\omega)KMAH}$$

KMAH index tracking through paraxial values

❖ 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 when crossing a caustic

Point source conditions

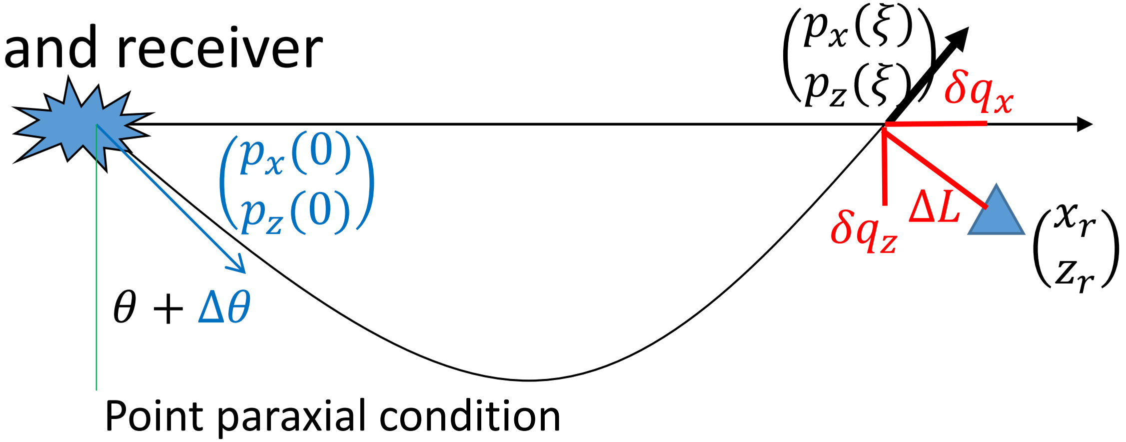
❖ In 3D, the determinant $\begin{vmatrix} p_x(\xi) & \delta q_x{}^4 & \delta q_y{}^4 & \delta q_z{}^4 \\ p_y(\xi) & \delta q_x{}^5 & \delta q_y{}^5 & \delta q_z{}^5 \\ p_z(\xi) & \delta q_x{}^6 & \delta q_y{}^6 & \delta q_z{}^6 \\ 0 & p_x(0) & p_y(0) & p_z(0) \end{vmatrix}$ may change sign.

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).

Two-points ray tracing with paraxial values

Consider the orthogonal distance between ray and receiver

$$\Delta L = (x_r - \delta q_x) p_z(\xi) - (z_r - \delta q_z) p_x(\xi)$$



Solve iteratively $\Delta L = \dfrac{dq_L}{d\theta} \; \Delta\theta$

or $\Delta L = \left(\dfrac{dq_x}{d\theta} p_z(\xi) - \dfrac{dq_z}{d\theta} p_x(\xi) \right) \Delta\theta$

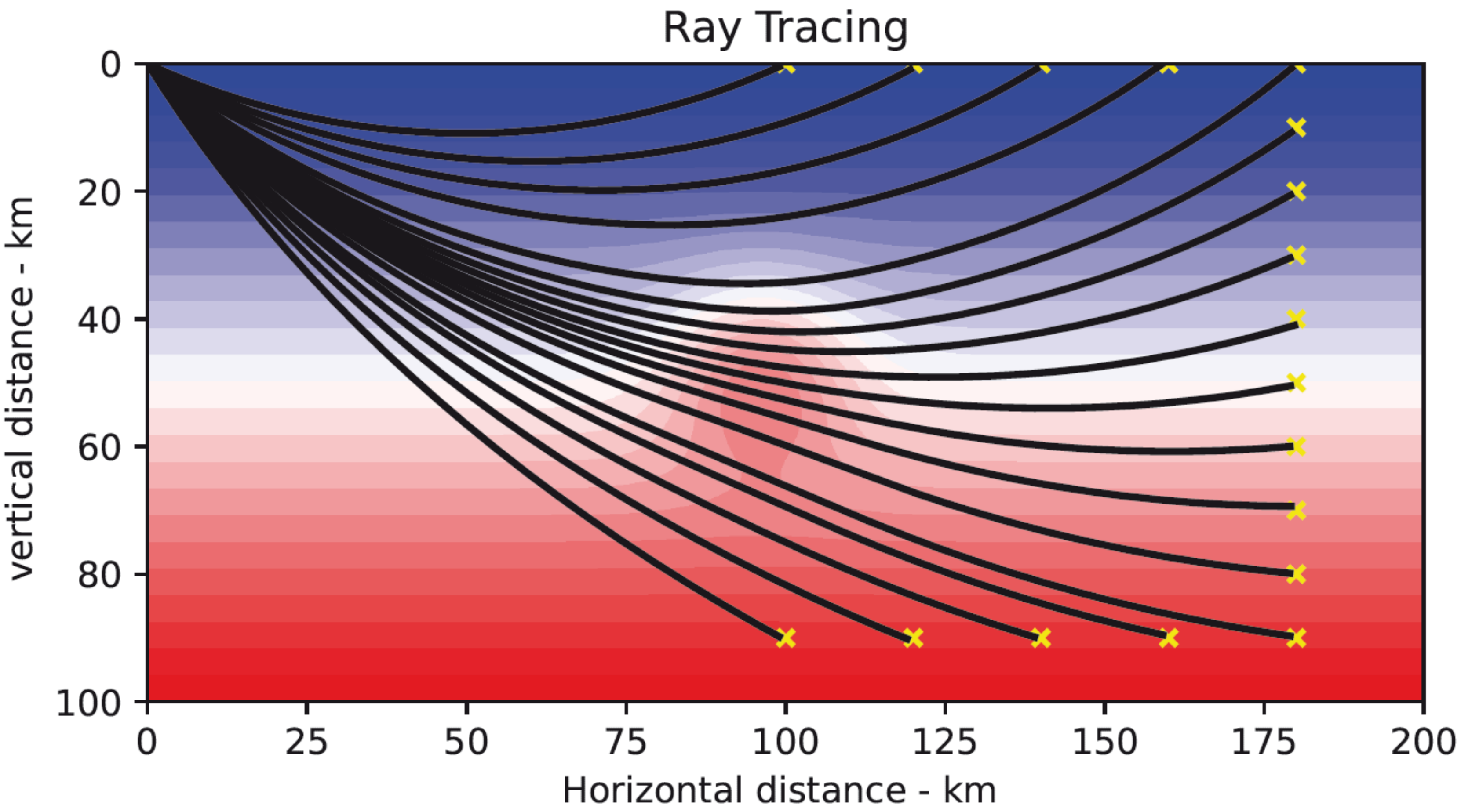
Derivative wrt shooting angle $\dfrac{dq_x}{d\theta} = \dfrac{\partial q_x}{\partial p_x} \dfrac{dp_x}{d\theta}(0) + \dfrac{\partial q_x}{\partial p_z} \dfrac{dp_z}{d\theta}(0)$ or $\dfrac{dq_x}{d\theta} = \dfrac{\partial q_x}{\partial p_x} p_z(0) - \dfrac{\partial q_x}{\partial p_z} p_x(0)$

$$\frac{dq_x}{d\theta} = \delta q_x 3 \;\; p_z(0) - \;\delta q_x 4 \;\; p_x(0) = \delta q_x$$

$$\frac{dq_z}{d\theta} = \delta q_z 3 \;\; p_z(0) - \;\delta q_z 4 \;\; p_x(0) = \delta q_z$$

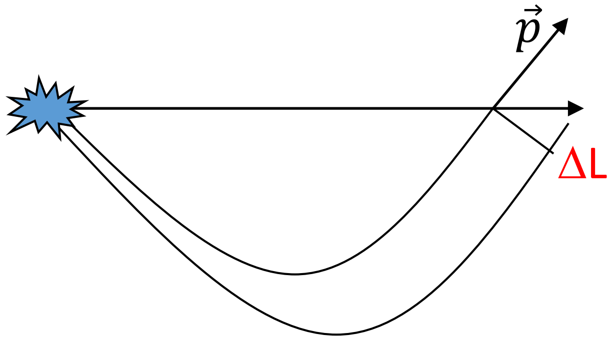
$$\Delta\theta = \frac{(x_r - \delta q_x) p_z(\xi) - (z_r - \delta q_z) p_x(\xi)}{\delta q_x p_z(\xi) - \delta q_z p_x(\xi)}$$

Two-points ray tracing



Amplitude estimation

Consider ΔL the distance between an exit point of a ray at the particule time ξ and the related paraxial ray.



From point paraxial ray δy^a

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

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

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

$$\frac{\Delta L}{\Delta\theta} = \frac{\begin{vmatrix} p_x(\xi) & \delta q_x 3(\xi) & \delta q_z 3(\xi) \\ p_z(\xi) & \delta q_x 4(\xi) & \delta q_z 4(\xi) \\ 0 & p_x(0) & p_z(0) \end{vmatrix}}{\sqrt{p_x(\xi)^2 + p_z(\xi)^2}}$$

Thanks to the point paraxial solutions , geometrical spreading $\Delta L/\Delta\theta$ can be computed, and the ray amplitude $A(\xi) \propto \frac{\Delta L}{\Delta\theta}$, and the KMAH index as well.

. *No need of solving the transport equation for getting the amplitude evolution*

Beam summation

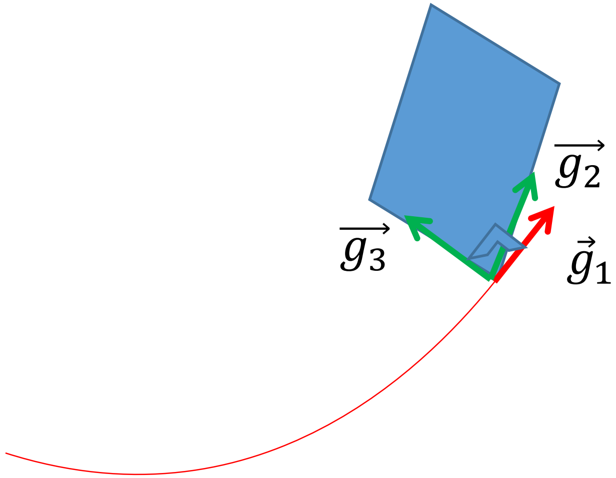
Using plane paraxial conditions with elementary solutions $\delta y1$ and $\delta y2$, we can construct any beam (as the Gaussian beam) in an ad-hoc summation …

In fact, one can combine any combination of elementary solutions $\delta y1, \delta y2$,$\delta y3, and\ \delta y4$ for focused beam summation (mixture of point and plane solutions)…

Remark: not too much used in seismology (except GBS), while extensive use in optics, such as the non-diffracting Airy solution (Lin et al, 2015)

Vibration: polarization estimation

Isotropic case: shear vibrations are orthogonal to compression vibrations



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

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

Time stepping

$$\frac{d\vec{g}_2}{dt} = \vec{g}_2.\nabla_{\vec{q}}c \;\; \vec{g}_1$$

$$\frac{d\vec{g}_3}{dt} = \vec{g}_3.\nabla_{\vec{q}}c \;\; \vec{g}_1$$

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

Particule stepping

$$\frac{d\vec{g}_2}{d\xi} = \frac{\vec{g}_2.\nabla_{\vec{q}}c}{p^2}\vec{g}_1$$

$$\frac{d\vec{g}_3}{d\xi} = \frac{\vec{g}_3.\nabla_{\vec{q}}c}{p^2}\vec{g}_1$$

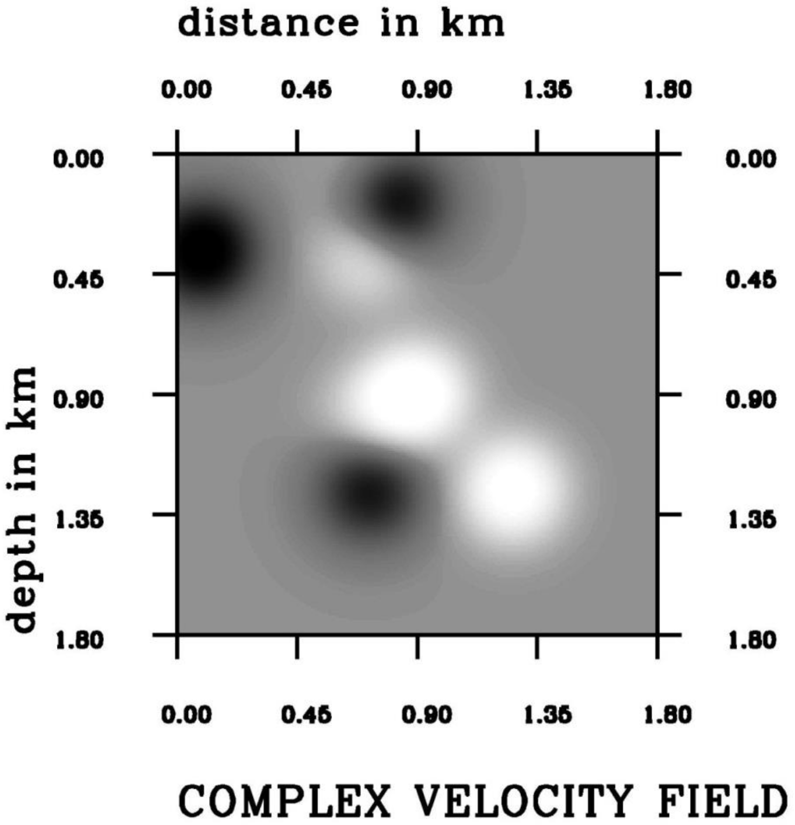
One additional equation for polarization

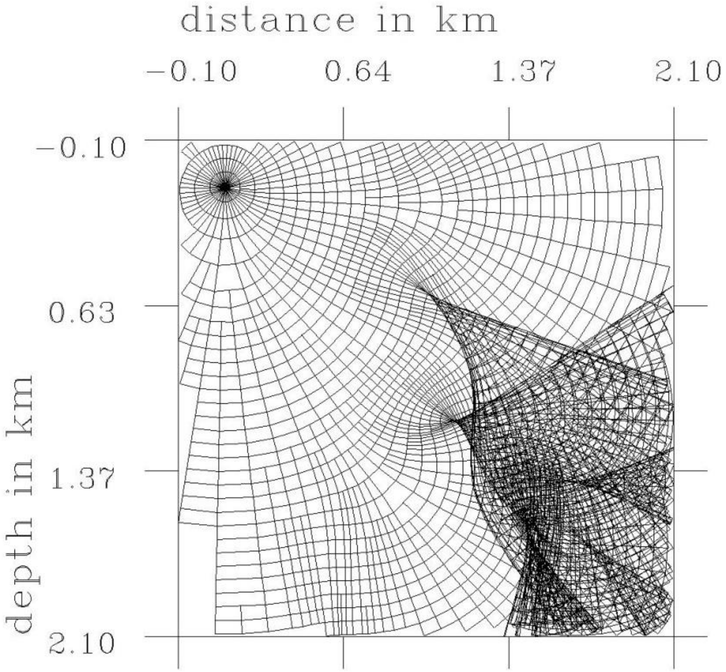
Chapman, 2004, p180

Step one: ray tracing

Example of a lagrangian-euler ray tracing

Lagrangian part: ray tracing
Eulerian part: wavefront sampling

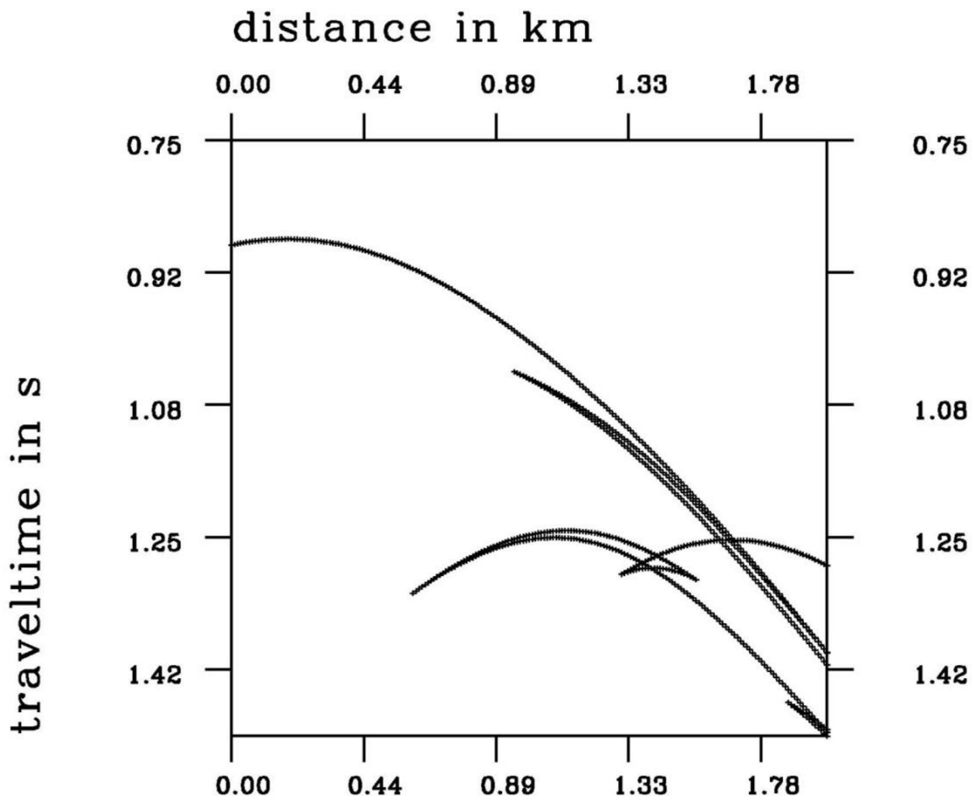




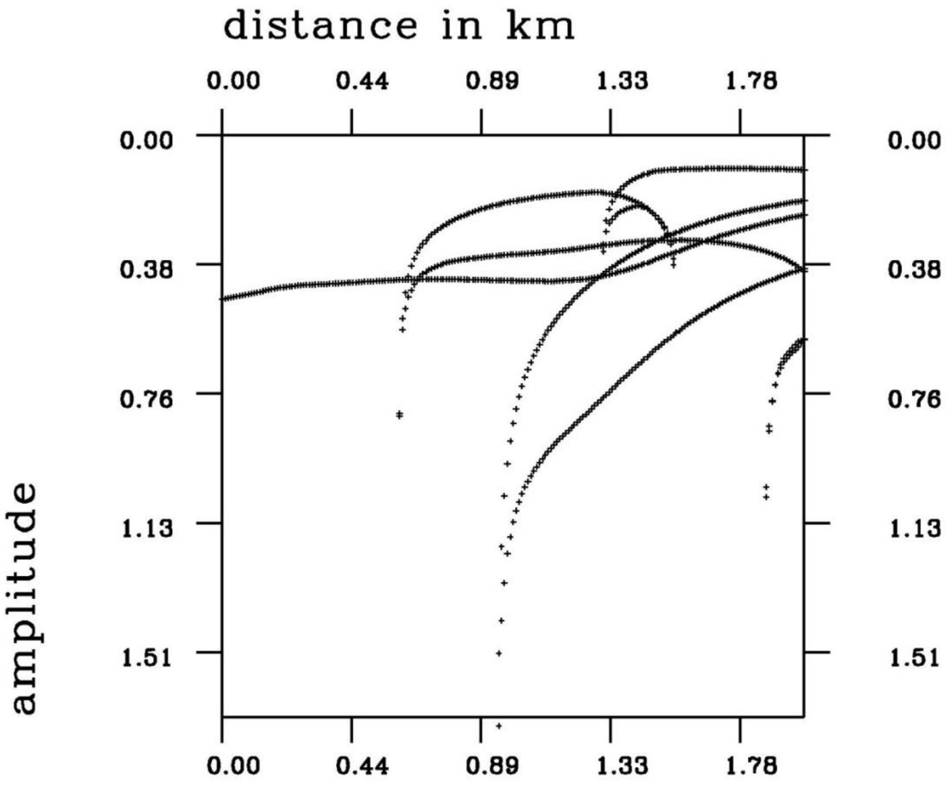
(Lambaré et al., 1996)

Step two: seismograms (paraxial)

(Lambaré et al., 1996)



TRAVELTIME X=2 km



AMPLITUDE X=2 km

Step three: polarization estimation (S waves)

Solving either $\vec{g}_2$ evolution or $\vec{g}_3$evolution, knowing that $\vec{g}_1$ is the direction of propagation.

The other polarization vector could be deduced from the orthonormal system $(\vec{g}_1, \vec{g}_2, \vec{g}_3)$.

Time stepping

$$\frac{d\vec{g}_2}{dt} = \vec{g}_2 . \nabla_{\vec{q}} c \;\; \vec{g}_1$$

$$\frac{d\vec{g}_3}{dt} = \vec{g}_3 . \nabla_{\vec{q}} c \;\; \vec{g}_1$$

Evolution of the polarization depends on the gradient of velocity $\nabla_{\vec{q}} c$

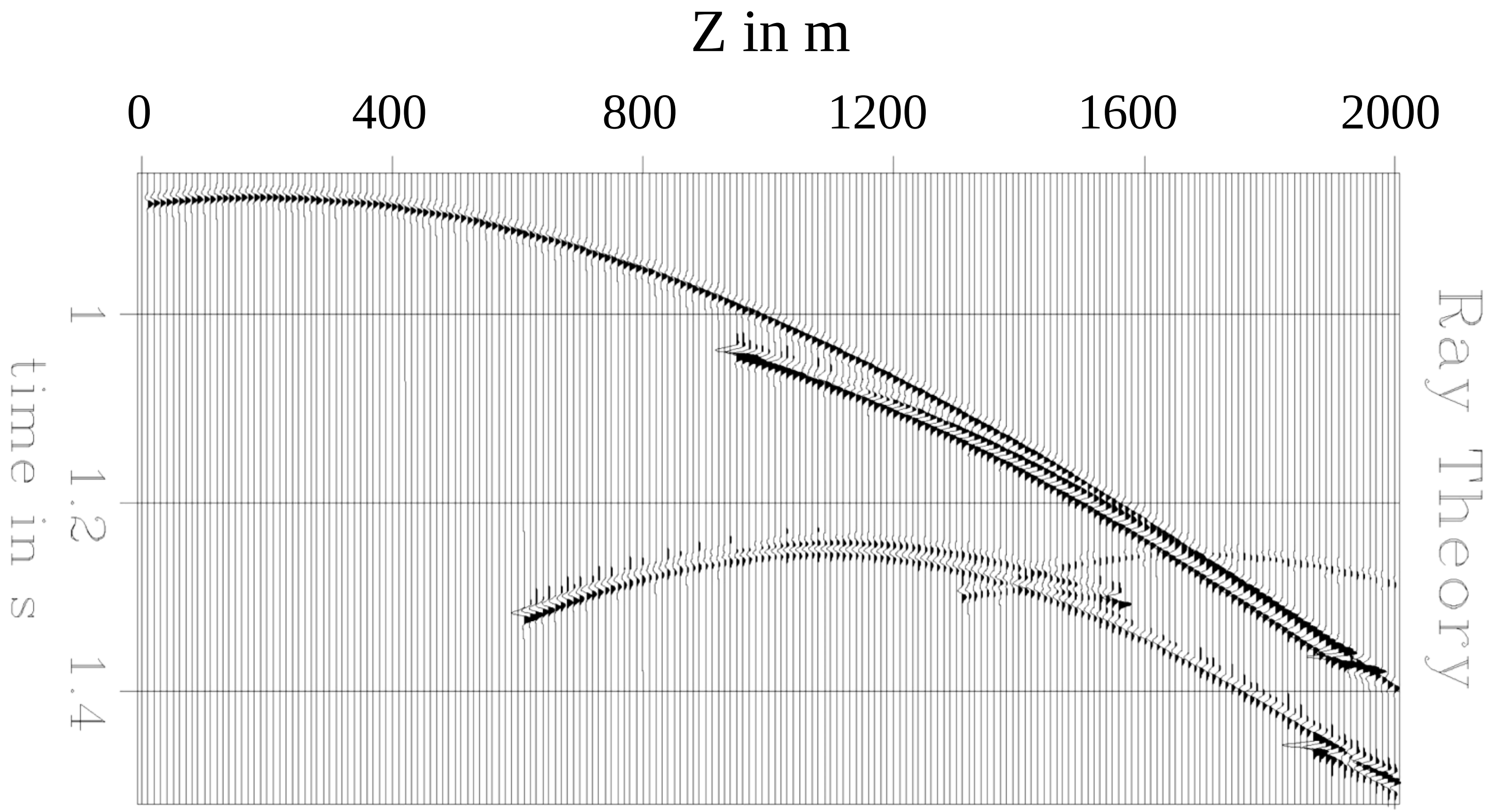
Particule stepping

$$\frac{d\vec{g}_2}{d\xi} = \frac{\vec{g}_2 . \nabla_{\vec{q}} c}{p^2} \vec{g}_1$$

$$\frac{d\vec{g}_3}{d\xi} = \frac{\vec{g}_3 . \nabla_{\vec{q}} c}{p^2} \vec{g}_1$$

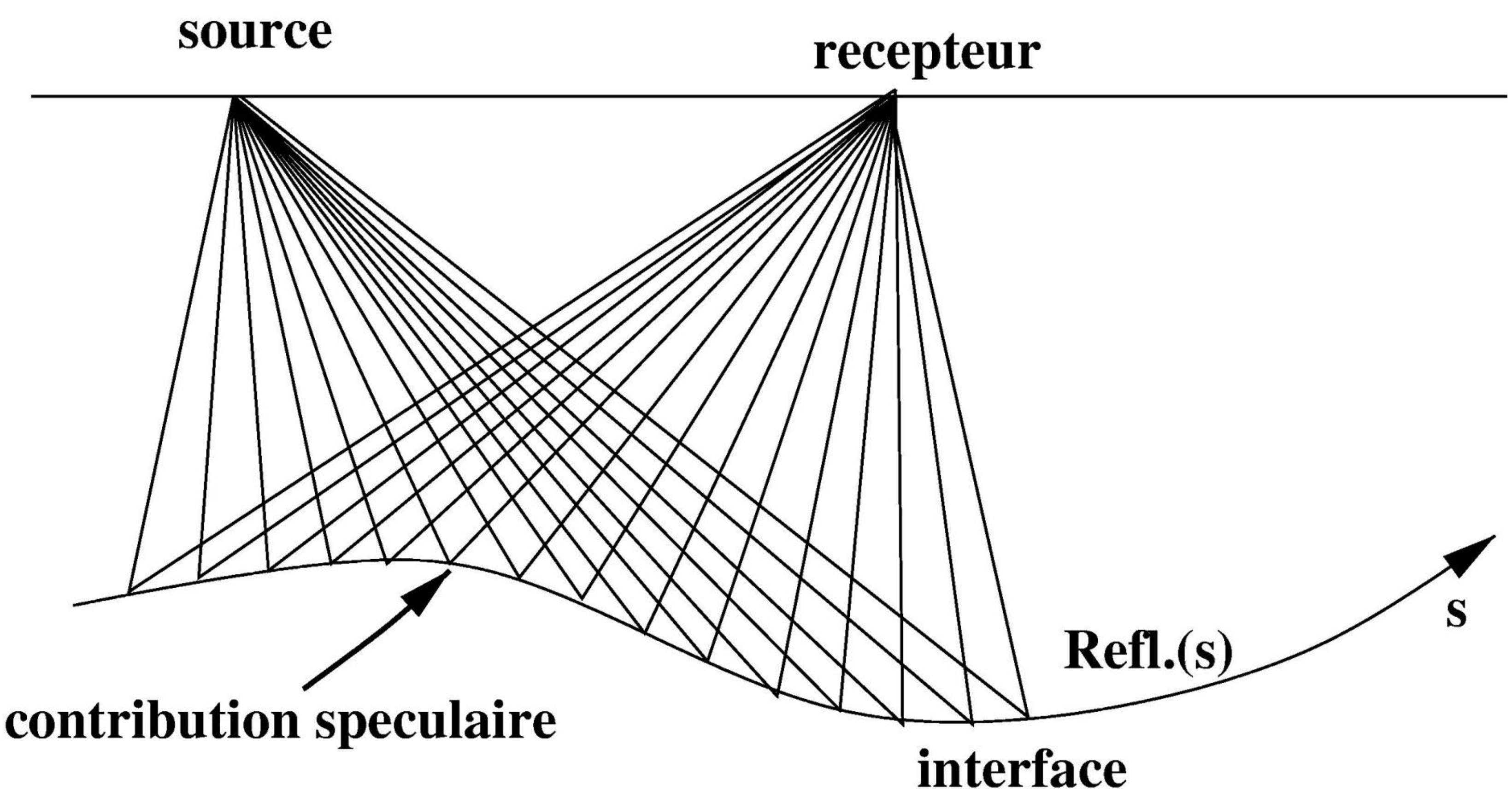
Useful for anisotropic investigation by
S receiver function theory
(Vinnik & Farra, 1992; Farra & Vinnik, 2000)

Step four: ray seismograms



(Lambaré et al., 1996)

Two-points ray tracing with boundary conditions



Interfaces in short!

Discontinuous model properties incompatible with ray ansatz:
two independent ray tracing on both sides of the interface to be connected

- ODE: resume solutions at the interface and restart the integration (Snell-Descartes law and paraxial Snell Descartes law)

(Cerveny et al, 1974; Farra et al, 1989)

- PDE: semi-lagrangian procedure back to ODE

(Rawlison & Sambdrige, 2003)

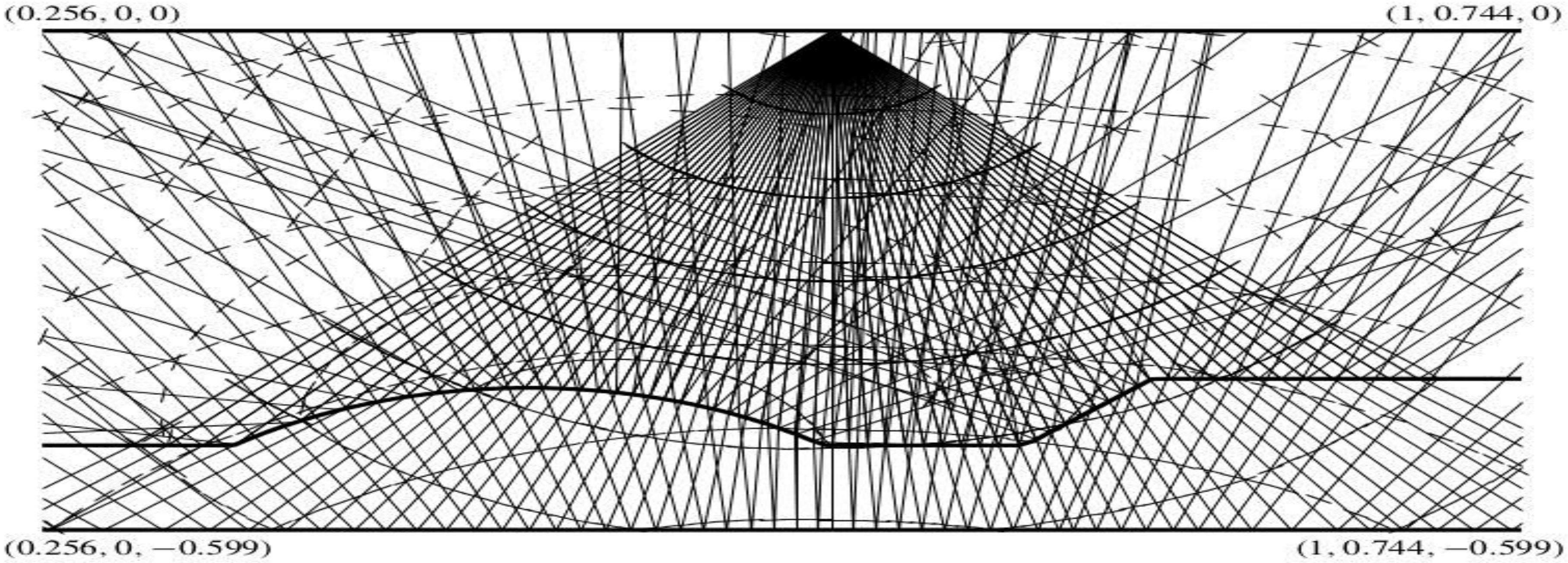
- PDE: rely on discontinuous methods

(Cheng & Shu, 2007)

Keeping complexity low?

More and more demanding on computer resources

3D French's model



Solutions: moving from ODEs to PDEs

(Osher et al, 2002) for adequate spatial sampling of the wavefront. Grids control the complexity!

Ray tracing: partial lessons to take away

Rays: a quite useful tool for interpretation and understanding

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

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

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

- Rays help the identification of phases: key interpretation
  - PDE does not allow easy interpretation!   (maybe work in progress …)

Toy computer codes for ray tracing

https://github.com/jeanvirieux/Tomography_training

Sub-directory: ray_tracing_analytic(.template)

comparison between analytic solutions and runge-kutta solutions

Sub-directory: ray-tracing_grid(.template)

ray tracing over a velocity grid using bspline interpolation

Sub-directory: two_points_ray_tracing(.template)

ray tracing including paraxial ray tracing with an illustration when hitting receiver

Simple codes based on python3 for practical understanding of the different equations of this presentation

Do it ...

Analyze python codes
and
run the shooting examples

Ray ODE vs Eikonal PDE

- **Scalar wave equation PDE**

  - Linear partial differential equation
  - Eulerian formulation

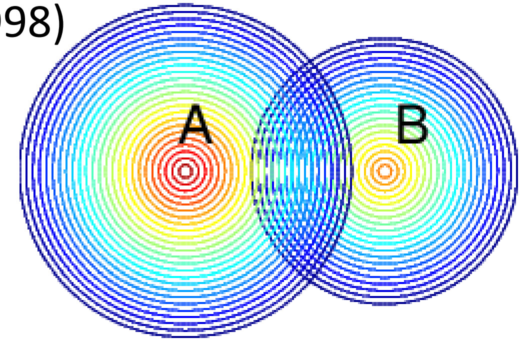
- **Ray ODE**: Methods of characteristics (Courant & Hilbert, 1966)

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

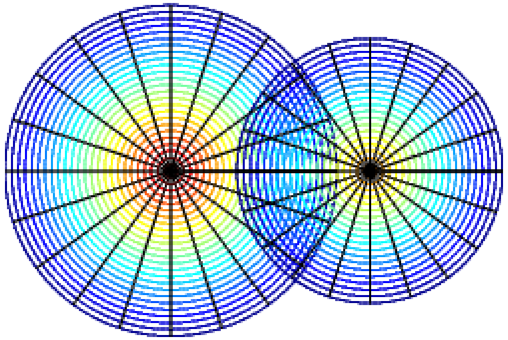
- **Eikonal PDE**: challenging equation (complete solution – the Graal)

  - Non-linear partial differential equations
  - Eulerian formulation as we compute quantities at fixed positions
  - Fastest solution through fast marching or fast sweeping methods

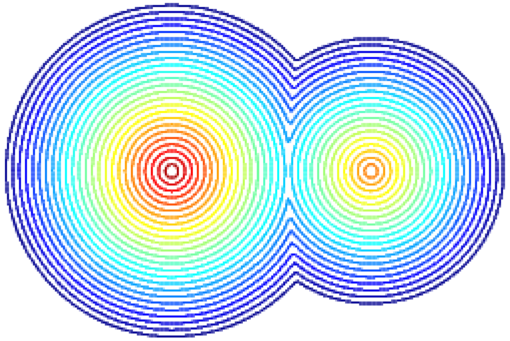
(after Runborg, 1998)



Wave solution



Ray solution



Eikonal solution

Asymptotic solution: beyond ray solution!

Geometrical theory of diffraction (Keller, 1962)

- Ray solution is **one** asymptotic solution among many other expansions.

  Ray Ansatz is limited to integer power of frequency

  Zero-out any diffraction effect (fractional frequency power)

- Airy, Bessel, Mathieu expansions (complicated formulations)!

  alternative expansions

  with or without diffraction

- <u>Fastest solution known as viscous solution (Crandall & Lions, 1983)</u>

Ray solution versus first-arrival solution

3.8.1 Ray Theory Travel Times and First-Arrival Travel Times

(Cerveny, 2001)

In this section, we shall define the ray-theory travel times and first-arrival travel times and explain the main differences between them.

One extracted property of first-arrival solution mentioned by Cerveny (2001)

c. The first-arrival travel time is a *unique* function of position. It is defined at any point of the model. There are no shadow zones. Moreover, the first-arrival travel time is a continuous function of coordinates. The first spatial derivatives of the first-arrival travel time, however, may be discontinuous. They may be discontinuous even at points where the velocities are continuous. (Example of such discontinuity include the intersection of the wavefronts of direct and head waves.)

Viscous solution!

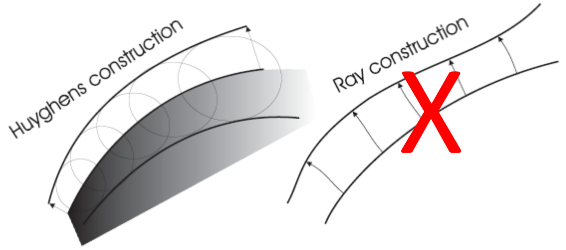
Understanding this asymptotic solution!

Defining its validity domain!

Crandall & Lions (1983)

Fermat/Huygens principles to Eikonal equation

First-arrival traveltimes follow Fermat principle of minimum time along any trajectory connecting the starting point and the end point. This principle is highly connected to the Huygens principle related to the wavefront construction



The related variational problem can be written

$$\delta \int u(s)ds = 0$$

$slowness\ u = {}^{1}/_{c}$

where $u(s)$ is the slowness at a given point and the curvilinear coordinate at this point is given by the quantity s.

The Eikonal equation is thought of the related PDE of this variational problem, leading to the Hamilton-Jacobi(-Bellman) equation.

(Kalaba, 1961 (isotropic case); Brandstatter, 1974 (anisotropic case)).

Fermat principle to Eikonal equation

In a 2D medium defined by coordinates (x, z), the path of the ray (perpendicular to the wavefront) is such that

$$\mathrm{T}(x,z) = \min_{l} \int u\big(x(l), z(l)\big) dl \qquad \textbf{(1)}$$

We consider an infinitesimal path Δl from $(x - sin\theta\ \Delta l,\ \mathrm{z} - cos\theta\ \Delta l)$ where the angle θ is the tangent angle to the current trajectory.

We get

$$\mathrm{T}(x,z) = \min_{\theta}[T(x - sin\theta\ \Delta l\ ,\ \mathrm{z} - cos\theta\ \Delta l) + u(x - sin\theta\ \Delta l\ ,\ \mathrm{z} - cos\theta\ \Delta l)\Delta l + \sigma(\Delta l)]$$

Expanding in Taylor series, we get

$$T(x,z) = \min_{\theta}\left[T(x,\mathrm{z}) - sin\theta\Delta l \frac{\partial T}{\partial x} - cos\theta \frac{\partial T}{\partial z} + u(x,z)\Delta l + \sigma(\Delta l)\right]$$

Or

$$\min_{\theta}[-sin\theta\ \Delta l\ T_x - cos\theta\ \Delta l\ T_z + u(x,z)\Delta l + \sigma(\Delta l)] = 0$$

When Δl is small, we get

$$sin\theta\ T_x + cos\theta\ T_z = u(x,z) \qquad \textbf{(2)}$$

From Lakshminarayanan and Varadharajan (1997)

with compact notation $T_x = \frac{\partial T}{\partial x}$ and $T_z = \frac{\partial T}{\partial z}$

Fermat principle to Eikonal equation

Minimizing with respect to θ gives

$$\sin\theta = T_x/\sqrt{T_x^2 + T_z^2} \text{ and } \cos\theta = T_z/\sqrt{T_x^2 + T_z^2}$$

Putting these expressions into equation (2) gives the eikonal equation (3)

$$T_x^2 + T_z^2 = u^2(x,z) \qquad \textbf{(3)}$$

This can be extended to 3D geometry as well

Fermat principle

$$\mathrm{T}(x,z) = \min_l \int u\big(x(l), z(l)\big)dl$$

Non linear Eikonal equation

$$T_x^2 + T_z^2 = u^2(x,z)$$

First-arrival time matches the zero-order ray time when it exists.
However, such time could be evaluated when there is no ray time.

No ray ansatz and frequency power expansion

From Lakshminarayanan and Varadharajan (1997)

Viscous solution versus ray solution

Fermat principle

$$\mathrm{T}(x,z) = \min_{s} \int u\big(x(s), z(s)\big) ds$$

Non-linear Eikonal equation

$$T_x^2 + T_z^2 = u^2(x,z)$$

also in 3D

First-arrival time matches the zero-order ray time when it exists.

However, such time could be obtained when there is no ray time.

No shadow zone!

No ray ansatz and no high-frequency asymptotic solutions.

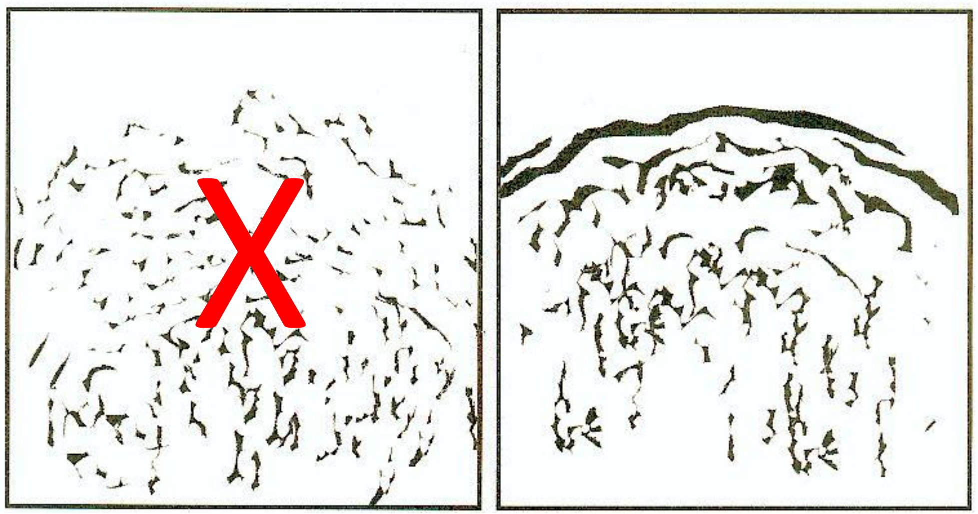
Still coherent wavefront!

Such Eikonal solution is sometimes called viscous solution.

Only one-value solution!

Multi-values « viscous » solution: the Graal!

Viscous solution with wavefront continuity



Viscous solution: first-arrival solution when wavefront continuity is preserved (maybe not differentiable!).

Diffraction is included: no shadow zone!

Asymptotic solution: beyond ray solution!

Viscous solution: efficient tools exist for computing it!

- Fast marching method O(N) for travel-times and for amplitudes
- Fast sweeping method O(N) for travel-times and for amplitudes
  - Amplitude equations have to be 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.

Viscous solution – first-arrival solution

A non-familiar interpretation of first phases

(often associated to HF approximation)

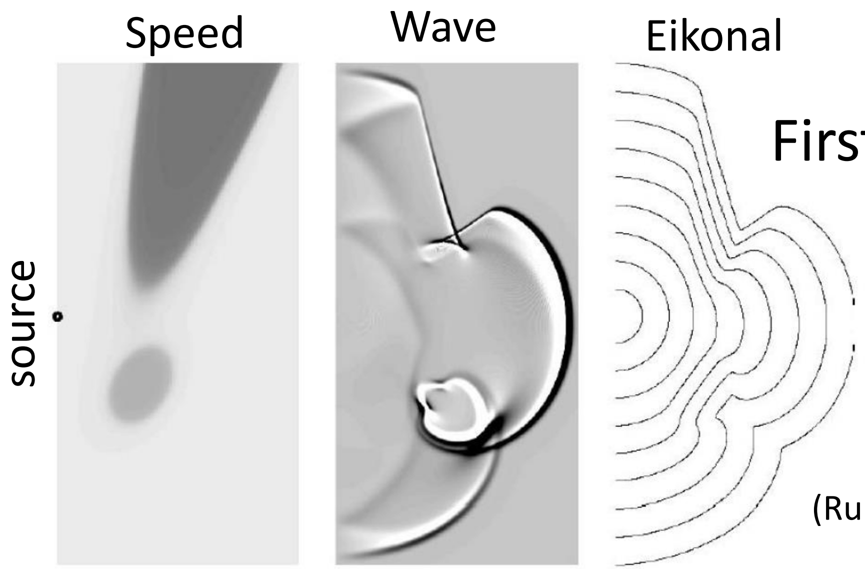
Wave disturbance (field discontinuity)

valid for any media

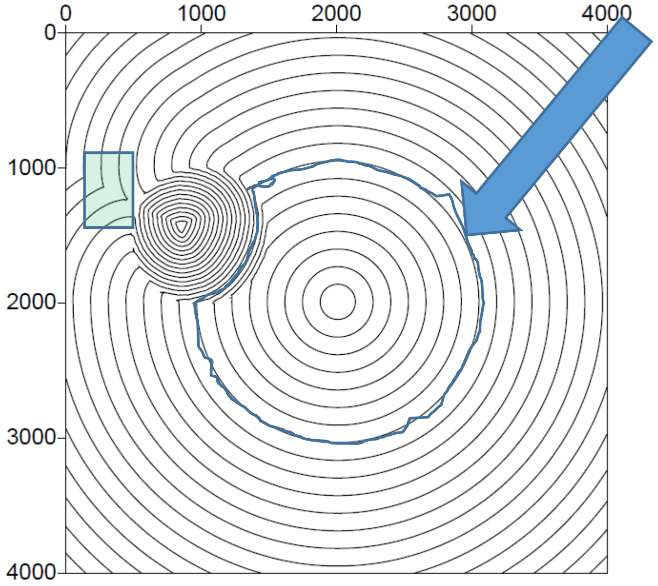
single value and always an answer

observable: continuous wavefronts

but possible discontinuous derivative



(Runborg, 2007)



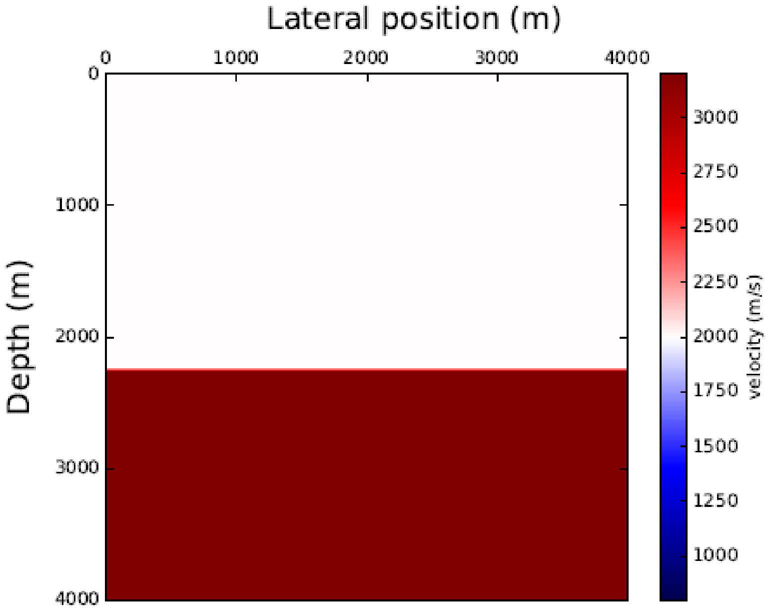
(Le Bouteiller, 2018)

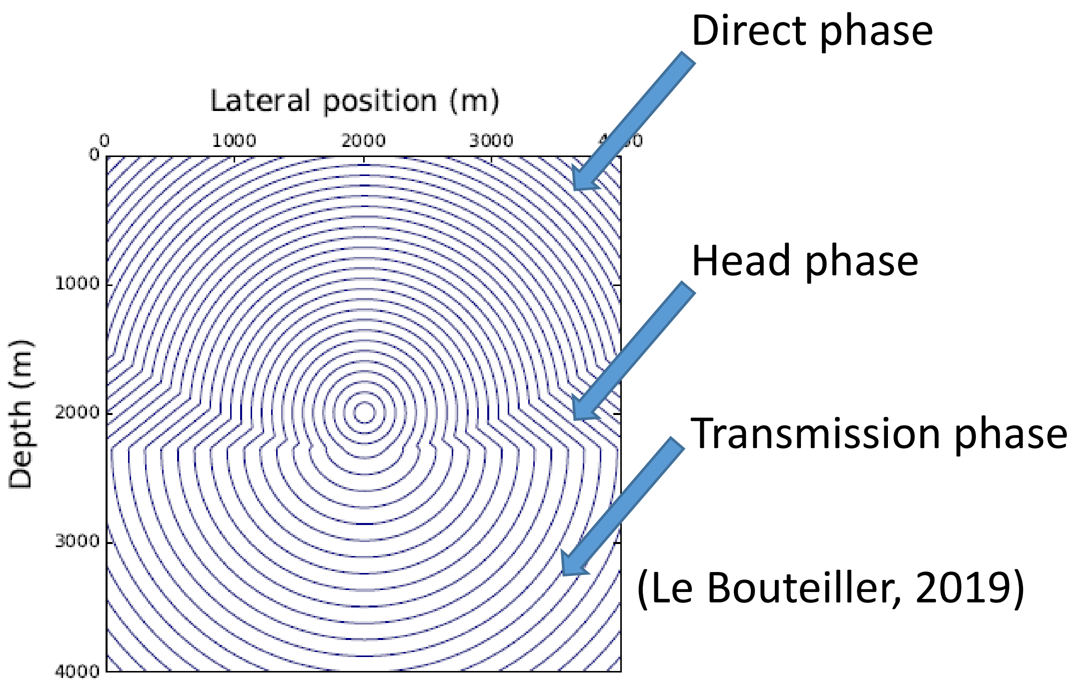
Line of same dansers

Wavefront: particles moving a synchronized way. They are in phase.

Diffraction effects included ($u \propto \omega^{1/2}$)

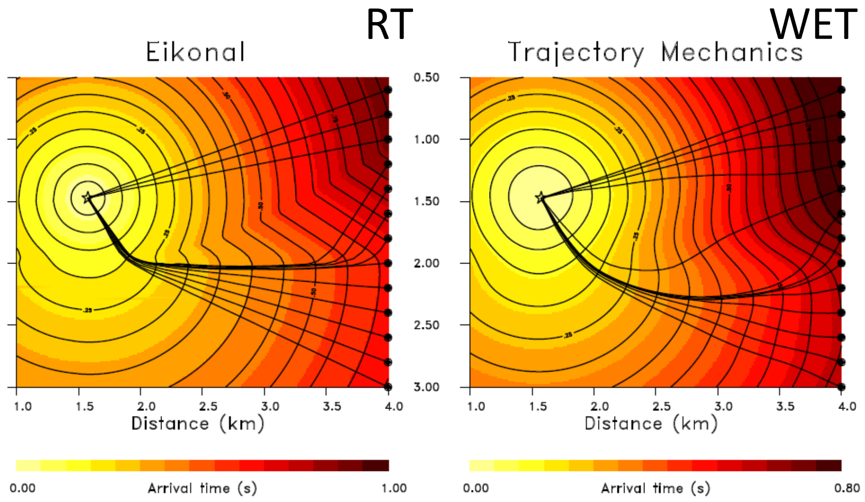
Example of continuous wavefront





(Le Bouteiller, 2019)

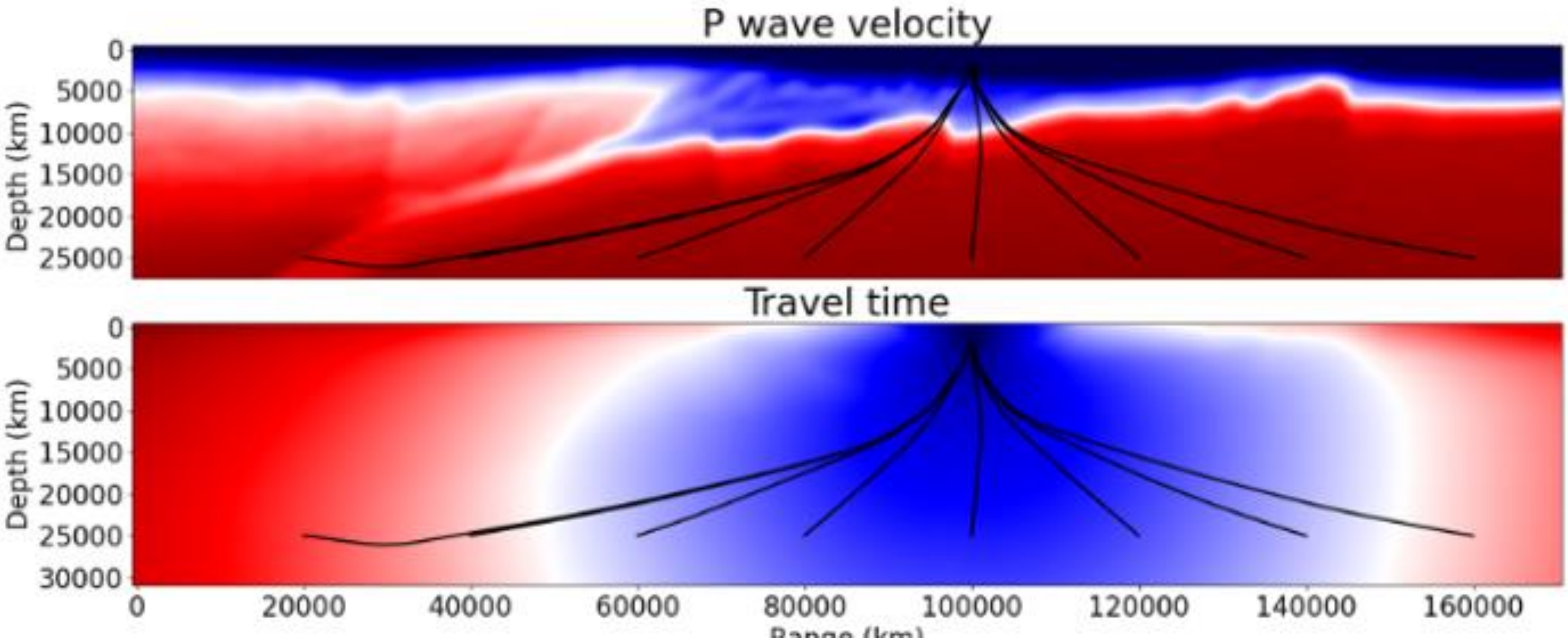
Meaning of paths when considering Eikonal solutions with sharp interfaces! ray or trajectory!



(Vasco & Nihei, 2019)

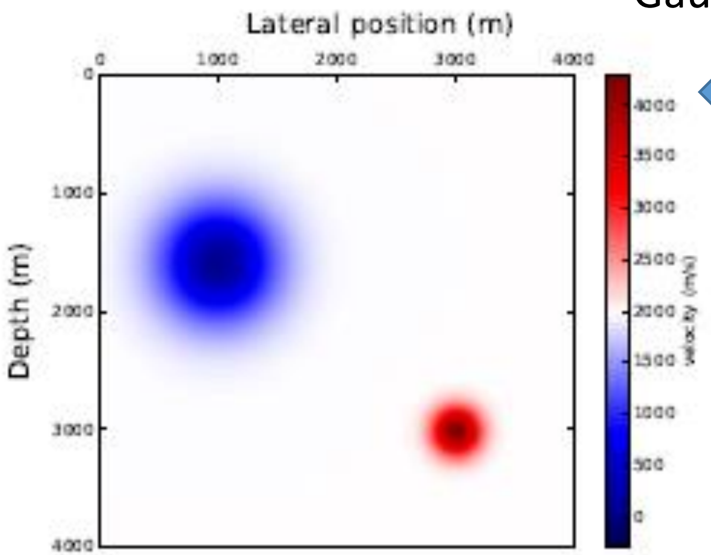
Viscous solution: minimum path …

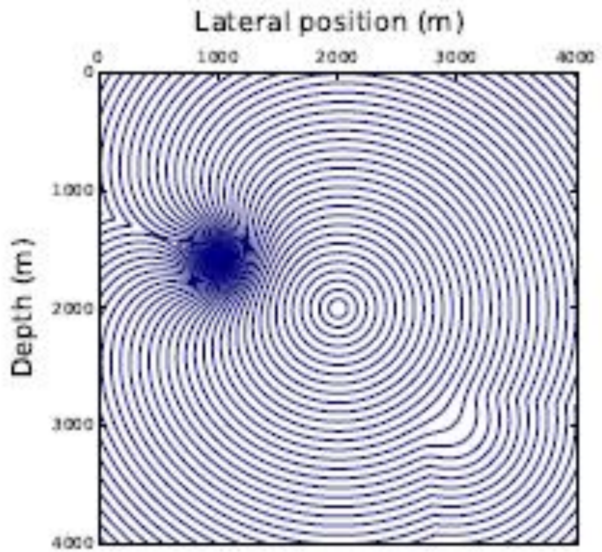
Minimal path between two points: Fermat principle



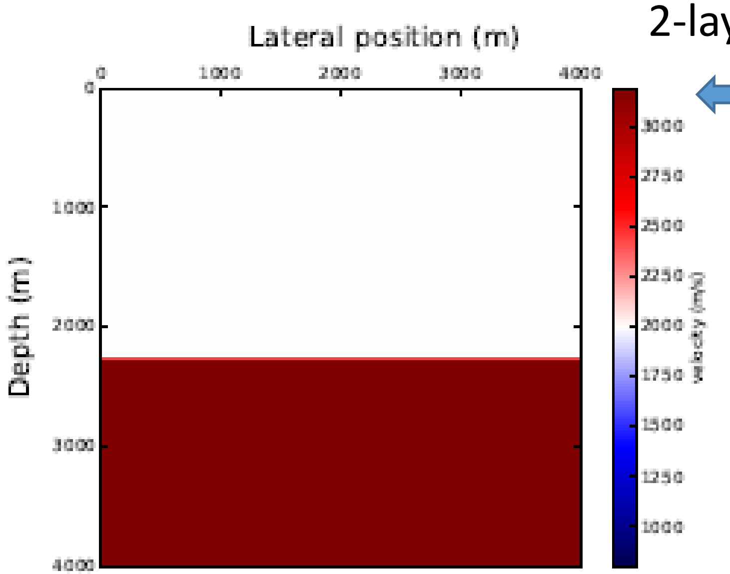
Viscous solution: examples (# ray solution)

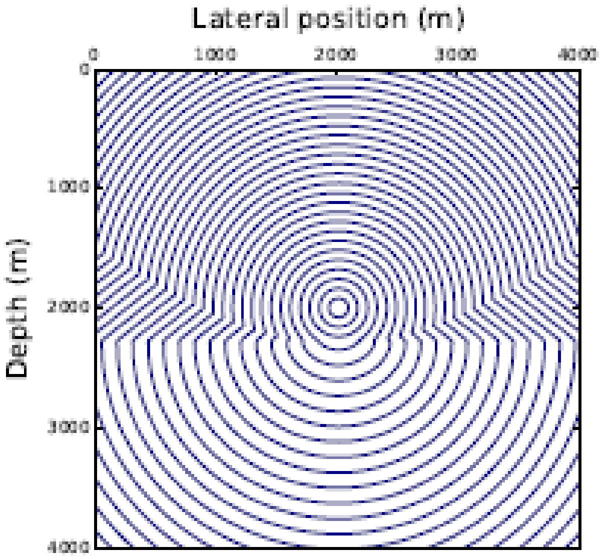
Gaussian model





2-layers model





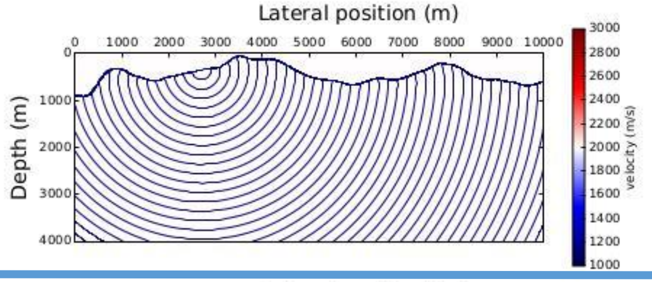
Viscous solution depends on the mesh discretization

(Le Bouteiller, 2018)

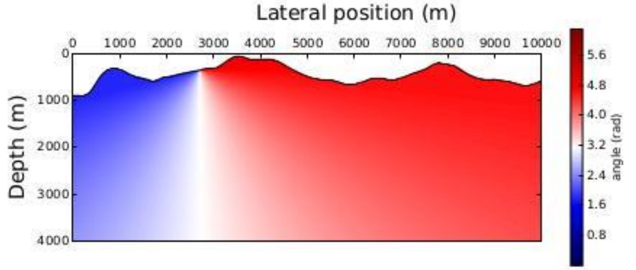
Transport equation and related PDEs

Isotropic case

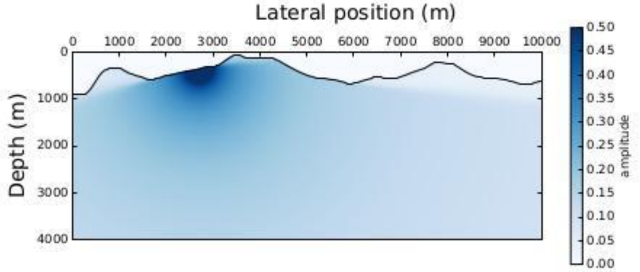
Non-linear PDE: Eikonal equation $(\nabla T)^2 - \frac{1}{c^2} = 0$



Linear PDE: take-off angle equation $\vec{\nabla}\varphi.\overrightarrow{\nabla T} = 0$

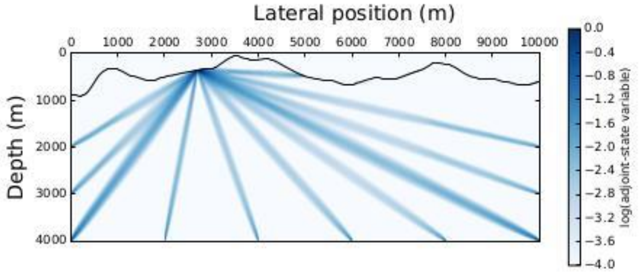


Linear PDE: amplitude equation $\vec{\nabla}.\left(A^2\overrightarrow{\nabla T}\right) = 0$



« transport without dissipation »

Linear PDE: adjoint equation $\vec{\nabla}.\left(\lambda\overrightarrow{\nabla T}\right) = \mathcal{F}$



Le Bouteiller (2018)

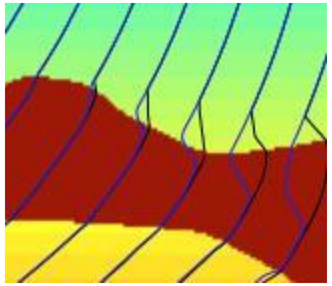
Useful family of equations

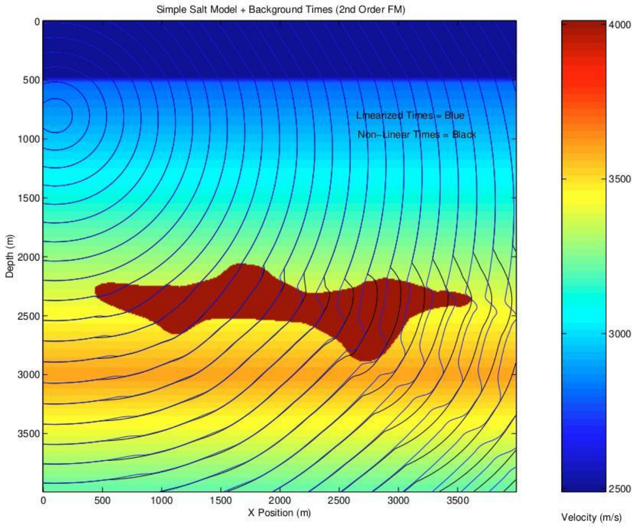
Linear Eikonal equations: time delay, angle, arclength estimation

Franklin & Harris (2001)

$$\overrightarrow{\nabla T}.\overrightarrow{\nabla \tau} = \mathcal{R}$$

Linear perturbation





Linear Transport equations: amplitude $\vec{\nabla}.\left(u\overrightarrow{\nabla T}\right) = 0$ singularities!

Belayouni (2013)

Factorization for removing these singularities at the source (and at receivers …)

$$\vec{\nabla}.\left(\bar{u}u_0T_0\overrightarrow{\nabla\tau}\right) + u_0\overrightarrow{\nabla T_0}.\overrightarrow{\nabla(\bar{u}\tau)} = 0$$

$$u = \bar{u}u_0; T = \tau T_0$$

with $known\ (analytical)\ solution\ \vec{\nabla}.\left(u_0\overrightarrow{\nabla T_0}\right) = 0$

Fast Marching Method : 3D geometry …

Computer codes available

Podvin & Lecomte (1996)

- Solving Eikonal equation for isotropic models: many codes

Cartesian and Spherical coordinates

- Solving Eikonal equation for anisotropic models: few codes

- Solving Eikonal equation and Sensitivity kernel: very few codes

Hamilton Fast Marching (HFM) and Adaptive Grid Discretizations (AGD) from Dr. Jean-Marie Mirebeau

HFM is written in C++17: Github repository (type Mirebeau and HFM on your browser)
Follow content of the file « Readme.md »
AGD is written in python & CUDA: Github repository (same location as the HFM software)
Recommendation of the installation from conda environment. See content of the file « Readme.md »

Fast Marching method: HFM illustration

Two interfaces for HFM library

Use of the FileHFM for any langage (C, Fortran …) at the expense of written files (not dramatic)

An interface with Python is included …

Attractive features in 2D geometry, 3D geometry, and on curved surface (Riemann metrics)

General anisotropy solver
Efficient TTI Eikonal solver,
Including topography through masks (or deformed Cartesian grid),
Computation of sensitivity kernels
Computation of adjoint field
Computation of rays

DO NOT WRITE YOUR OWN CODE! (**Pykonal, pyekfmm, scikit-fmm** from github … among many other codes)

Ray and Eikonal toy examples

https://github.com/jeanvirieux/Tomography_training

See the README.md

Do it …

Analyze python codes

and

Compile HFM

Run simple examples

Take-away message

- ❑ Ray solution (multi-valued)

  When available, fruitfull for interpretation

- ❑ Viscous solution (single-valued)

  Efficient computer codes, even for anisotropy (TTI)

  No ray perturbation for anisotropy …

- ❑ Viscous solution (multi-valued)?

  Still open problem for efficient numerical strategy

  Such solution is single-valued in the phase space!!!

Forward problem: Temporary End

End of the first part

New search direction:

Neural Eikonal Solver: physics-informed Neural Network

PINNeikonal from github … (U. Waheed)

peikonal from github … (J. Calder)

Cool feature: open road for efficient tomography strategies …

(work in progress … see relation with the second part of this presentation)

Outline on first-arrival traveltime tomography

- Images at very different scales
- Waves and Phases: various concepts
- Few points on first-break ray-based tomography
- Illustration on 30-years Western Alps tomography
- First-break eikonal-based tomography
- First-break wave-equation-based tomography
- Hypocenter-velocity joint inversion
- Conclusion

Ray-based tomography: Fréchet derivative

Delayed ray-based tomography based on Fréchet derivative building

Delayed traveltime tomography

$$t(source, receiver) = \int_s^r s(x,y,z)dl$$

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

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

related to linear Eikonal equation

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

$$t(s,r) \approx \int_{s_0}^{r_0} s_0(x,y,z)dl + \int_{s_0}^{r_0} \delta s(x,y,z)dl$$

« frozen » ray approximation (ray connecting source/receiver for the known slowness s_0)

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

No Fermat argumentation!

$$\delta t(s,r) \approx \int_{s_0}^{r_0} \delta s(x,y,z)dl$$

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.

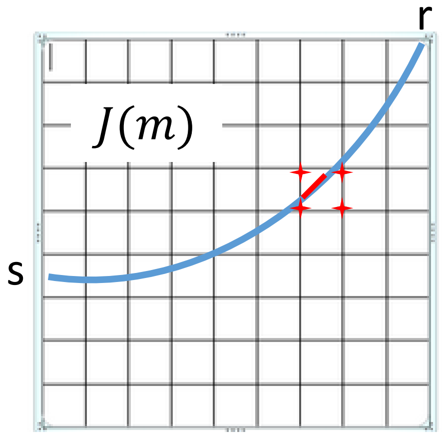
For each node, we specify a value $\delta s_{i,j,k}$. The interpolation will be performed with functions as step functions. For each grid point (i,j,k), shape functions $h_{i,j,k} = 1$ for i,j,k, and zero for other indices.

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

Other shape functions are possible with two-end members (nodal versus modal):
fourier functions (cos,sin), chebychev, spline, wavelet … and so on

Sampling the model space is the mandatory stabilization strategy
(smoothing or damping ones)
Model discretization provide an implicit limit to the wavenumber range to be filled in

Discrete linearized inversion problem



Discretization of the model fats the ray

$$\delta t(s,\mathrm{r}) = \int_{ray_0} \sum_{cube} \delta s_{i,j,k}\, h_{i,j,k}\, dl = \sum_{cube} \delta s_{i,j,k} \boxed{\int_{ray_0} h_{i,j,k} dl}$$

Weighted ray segment

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

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

$$\delta t(n) = J(n,m)\, \delta s(m)$$

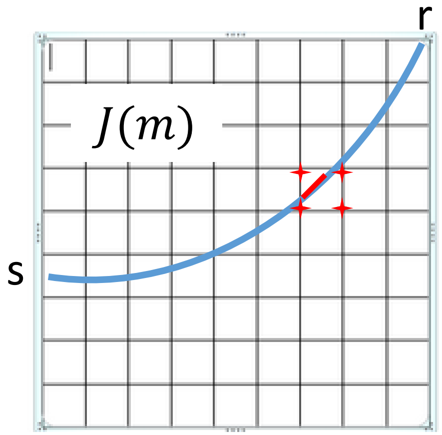
to be solved in least-squares sense

$$\begin{pmatrix} \delta t_1 \\ \delta t_2 \\ \vdots \\ \delta t_{n-1} \\ \delta t_n \end{pmatrix} = \begin{pmatrix} \frac{\partial t_1}{\partial s_1} & \cdots & \frac{\partial t_1}{\partial s_m} \\ \frac{\partial t_2}{\partial s_1} & \cdots & \frac{\partial t_2}{\partial s_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial t_{n-1}}{\partial s_1} & \cdots & \frac{\partial t_{n-1}}{\partial s_m} \\ \frac{\partial t_n}{\partial s_1} & \cdots & \frac{\partial t_n}{\partial s_m} \end{pmatrix} \begin{pmatrix} \delta s_1 \\ \delta s_2 \\ \vdots \\ \delta s_{m-1} \\ \delta s_m \end{pmatrix}$$

Sensitivity matrix J is a sparse matrix

also named Fréchet derivative or Jacobian matrix …

Discrete linearized inversion problem



Discretization of the model fats the ray

$$\delta t(s,r) = \sum_{i,j,k} \delta s_{i,j,k} \Delta l_{i,j,k}$$

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

(billion, million)

$$\delta t(n) = J(n,m)\, \delta s(m)$$

- n dimension of the data space
- m dimension of the model space

$$\begin{pmatrix} \delta t_1 \\ \delta t_2 \\ \vdots \\ \delta t_{n-1} \\ \delta t_n \end{pmatrix} = \begin{pmatrix} \frac{\partial t_1}{\partial s_1} & \cdots & \frac{\partial t_1}{\partial s_m} \\ \frac{\partial t_2}{\partial s_1} & \cdots & \frac{\partial t_2}{\partial s_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial t_{n-1}}{\partial s_1} & \cdots & \frac{\partial t_{n-1}}{\partial s_m} \\ \frac{\partial t_n}{\partial s_1} & \cdots & \frac{\partial t_n}{\partial s_m} \end{pmatrix} \begin{pmatrix} \delta s_1 \\ \delta s_2 \\ \vdots \\ \delta s_{m-1} \\ \delta s_m \end{pmatrix}$$

Least-squares solution

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.

Least-squares solution

$$J^t J \delta s = J^t \delta t$$

$$\delta s = (J^t J)^{-1} J^t \delta t$$

Remark

- Another interesting solution with minimum norm

Least-norm solution

$$J J^t \delta s = \delta t$$

$$\delta s = J^t (J J^t)^{-1} \delta t$$

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

Damped least-squares solution

$$\delta t = J\delta s$$
$$d = Gm$$
$$b = Ax$$

Damping parameter ε

A is a rectangular matrix (either over- or under-determined)

$$\min_x ||Ax - b||^2 + \varepsilon\, ||x||^2$$

LSQR solves it using only products Ax or $A^T b$ by considering the system

$$(A^T A + \varepsilon\, I)x = A^T b$$

widely used subroutine in travel-time tomography

LSMR solves it using only products Ax or $A^T b$ by considering the system

$$(A^T A + \varepsilon\, I)x = A^T b$$

http://www.numerical.rl.ac.uk/spral/doc/latest/Fortran/

Occam’s razor

➢ Do not use more complicated maths than the data deserves

➢ Approximate the least constrained quantity

Given: data (observed and modeled)
Assumed: wavefront 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.

Ockham (~1295-~1349)

Constable et al (1987)

Constrained damped least-squares solution

Constrained damped least-squares solution

$$\min_x (||Ax - b||^2 + \lambda ||Dx||^2 + \varepsilon \, ||x||^2)$$

D operator is a smoothing operator, such as a Laplacian operator which limit variations of the spatial second derivative of the slowness model.

Two hyper-parameters λ and ε to be selected?

Penalty approach is often selected

$$\begin{bmatrix} \mathcal{g}_k \\ \varepsilon I \end{bmatrix} [\delta s_k] = \begin{bmatrix} \dfrac{\partial t}{\partial s_k} \\ \lambda D \\ \varepsilon I \end{bmatrix} [\delta s_k] = \begin{bmatrix} \delta t_k \\ 0 \\ 0 \end{bmatrix}$$

Smoothing could vary with coordinates

$$\lambda_x D_x + \lambda_y D_y + \lambda_z D_z$$

with seven-points finite-difference stencil along each direction for the laplacian

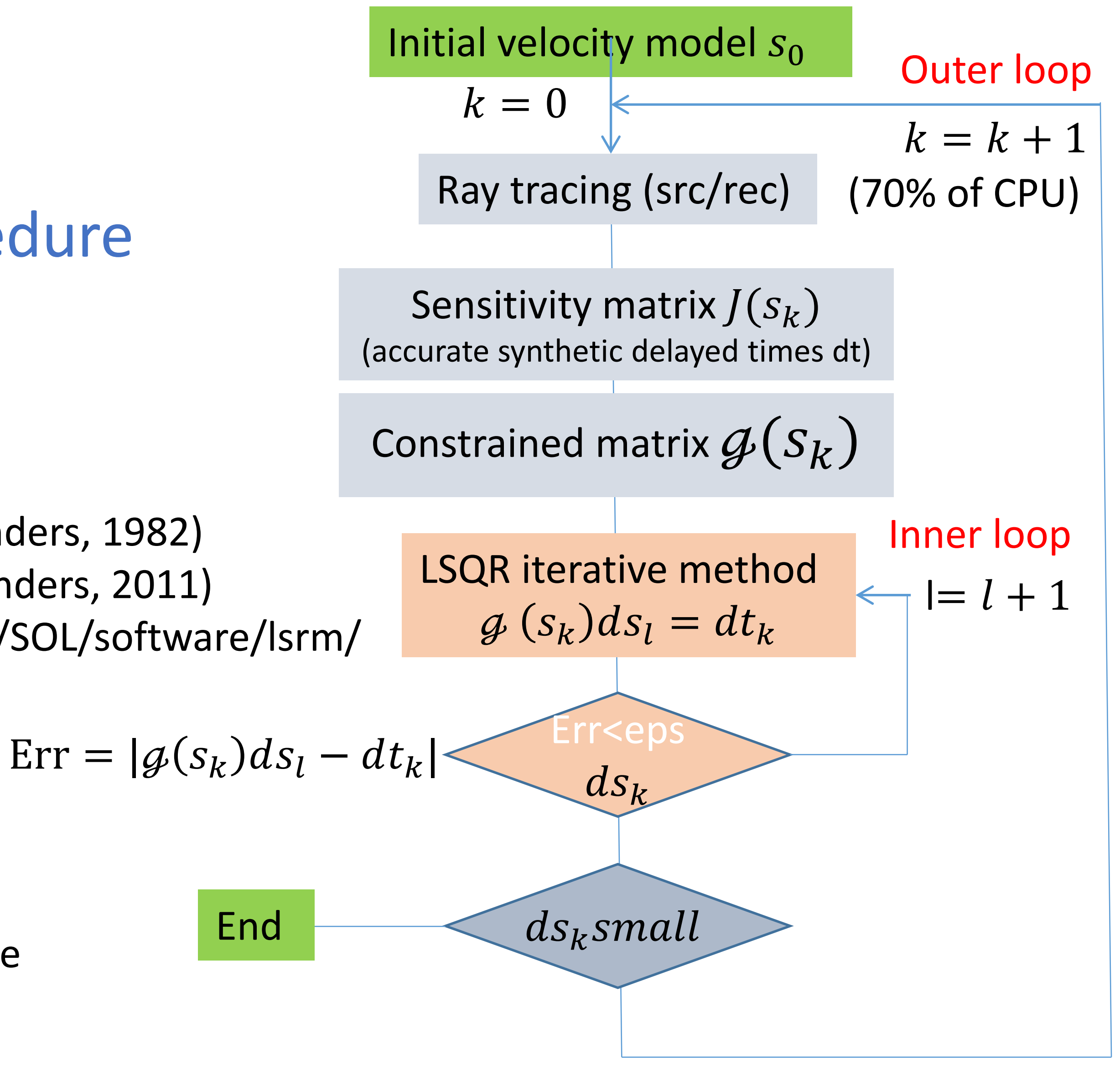
Hyper-parameter λ non-zero in all codes

Discrete Fréchet-ray algorithm

Two-loops procedure

LSQR subroutine (Paige and Sanders, 1982)
LSRM subroutine (Fong and Sanders, 2011)
http://web.stanford.edu/group/SOL/software/lsrm/

LSQR or LSRM includes the damping operator

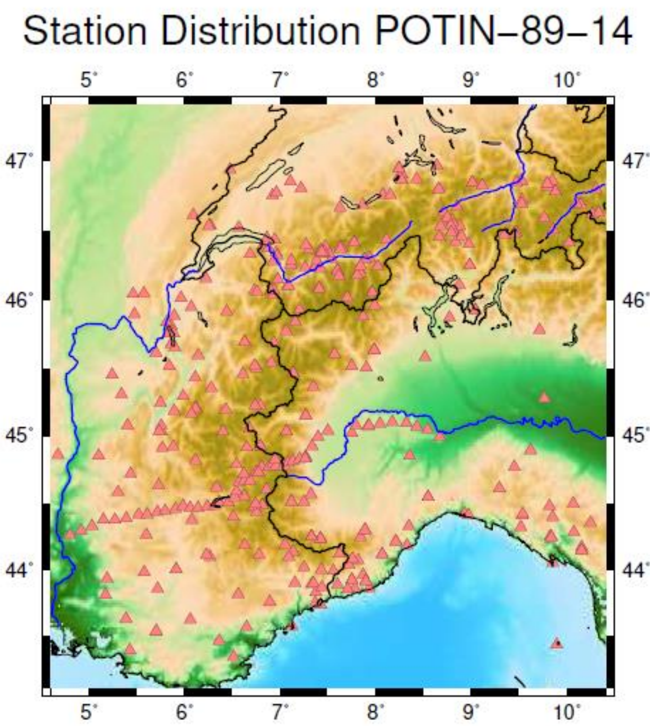


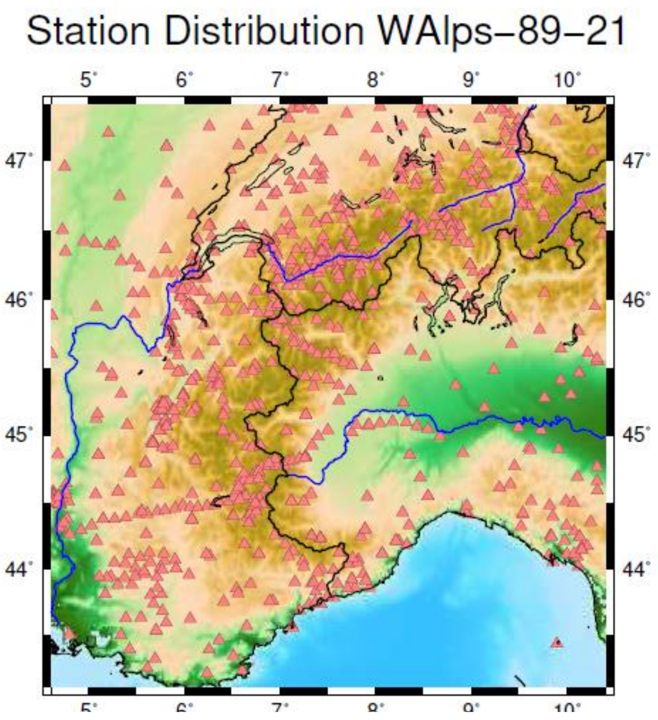
The slowness field is denoted by s and it is often the one we reconstruct

Outline on first-arrival traveltime tomography

- Images at very different scales
- Waves and Phases: various concepts
- Few points on first-break ray-based tomography
- Illustration on 30-years Western Alps tomography
- First-break eikonal-based tomography
- First-break wave-equation-based tomography
- Hypocenter-velocity joint inversion
- Conclusion

SISMALP: 30 years Western Alps recording





Increasing number of permanent stations

Starting with manual picks performed by B. Potin: a human-eye checking

Downloading picks from ReNass (FR), RSNI/DipTeris (IT) through ISC and SED (CH)

Automatic cleaning of picks by matching predicted picks from HQ-89-14 model in a 10s interval.

Not a tectonic interpretation on my own: I am not able to do so

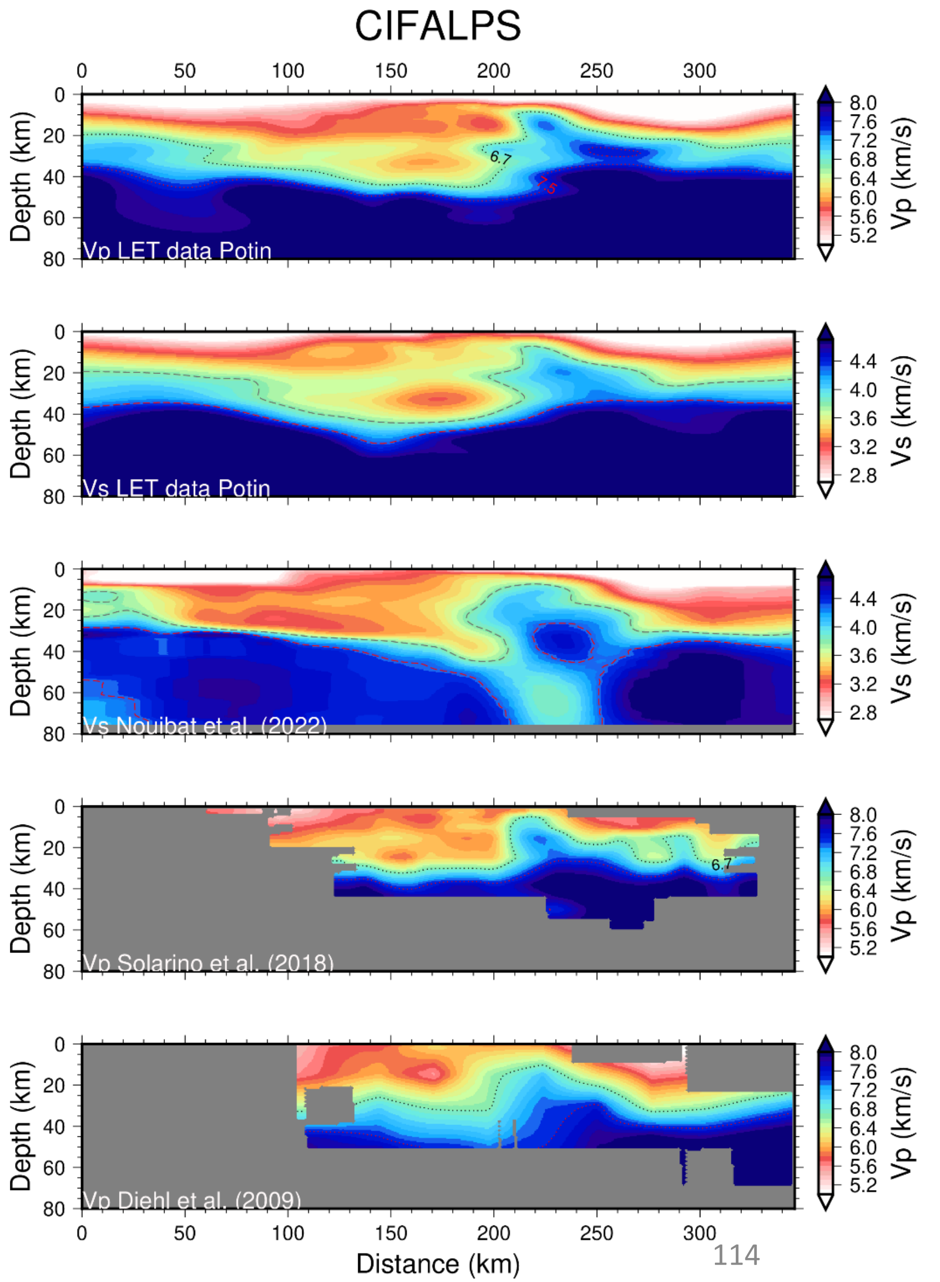
Such a target zone has been investigated by many groups using different methods

Database	Selection	Events	Stations	P picks	S picks
POTIN-89-14	All picks	54 409	373	542 818	460 129
HQ-89-14	> 12 P & 6 S	13 022	367	309 228	263 498
SQ-89-14	> 6 picks	50 331	373	533 499	451 517
WAlps-89-21	All picks	82 088	1043	952 317	670 786
WA-89-21	> 6 picks	75 538	1043	936 977	661 522

Western Alps: a target zone

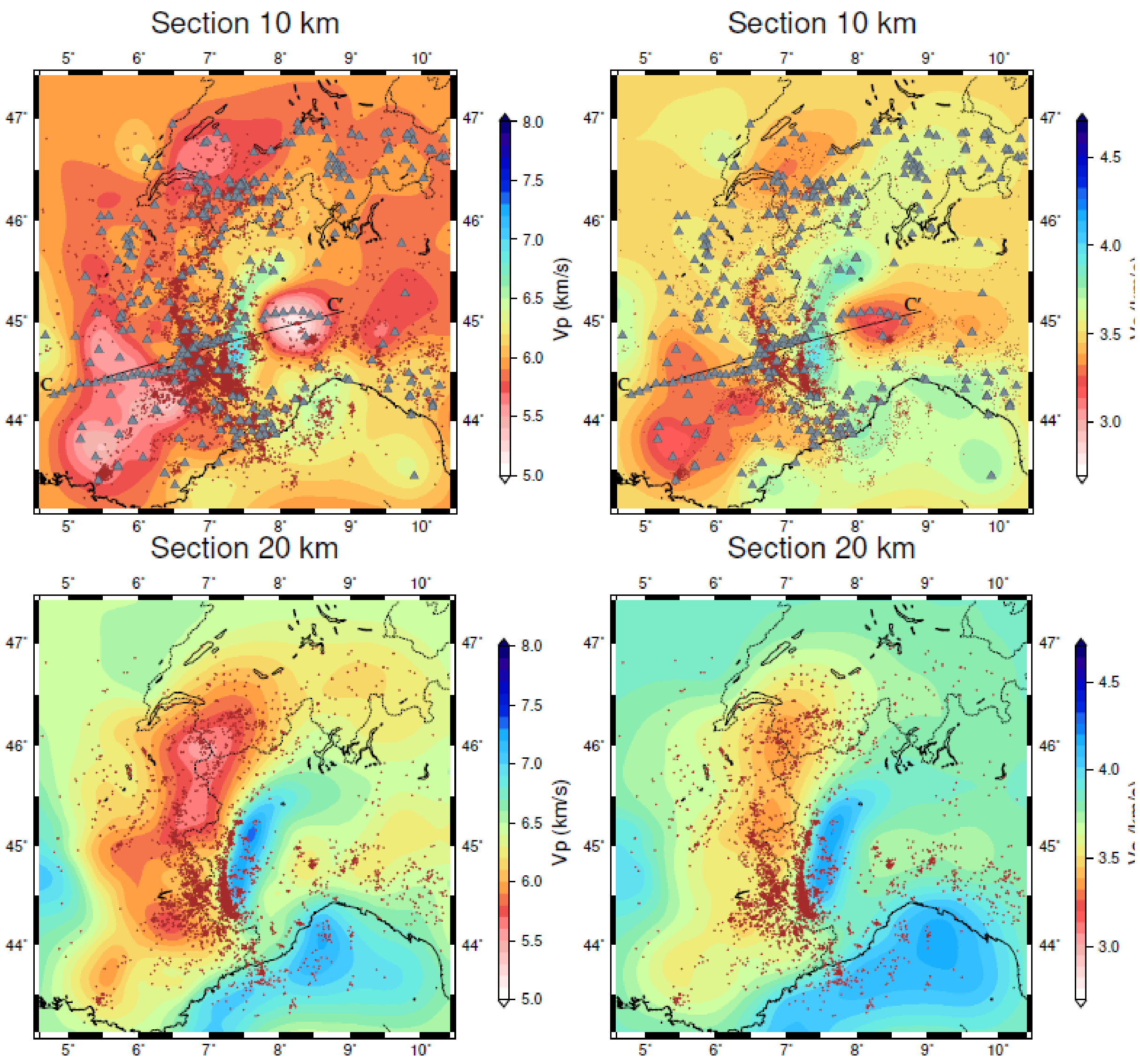
- Local earthquake tomography (Solarino et al., 1997; Paul et al., 2001; Diehl et al., 2009a,b; Solarino et al., 2018)
- Controlled source investigation (ECORS-CROP Deep Seismic Sounding Group, 1989; Nicolas et al., 1990; Thouvenot et al., 2007)
- TransD(?) ambient-noise tomography (Stehly et al., 2009; Lu et al., 2018; Zhao et al., 2020; Nouibat et al., 2022)
- Receiver-function approach (Zhao et al., 2015; Paul et al., 2022)
- Teleseismic traveltime tomography (Zhao et al., 2016; Paffrath et al., 2021)
- Teleseismic full waveform inversion (Beller et al., 2018)

Focus on methodological impacts
See the nice paper by Malusà et al (2021) for geodynamic interpretation

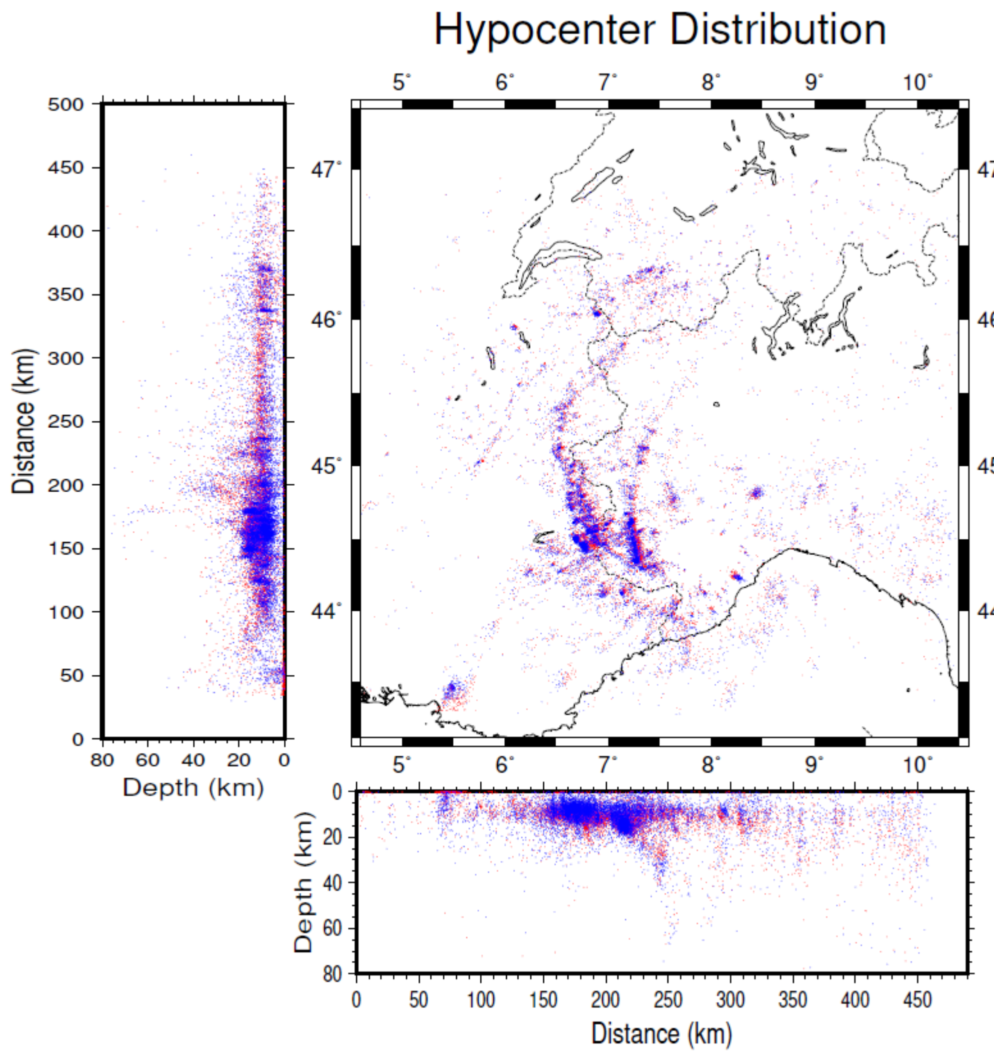


(from A. Paul)

Vp & Vs models: penalty approach

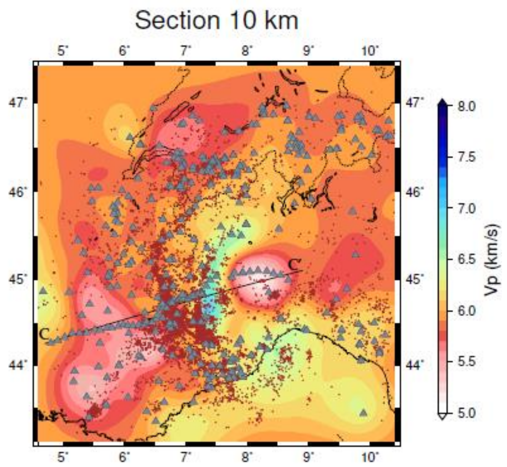


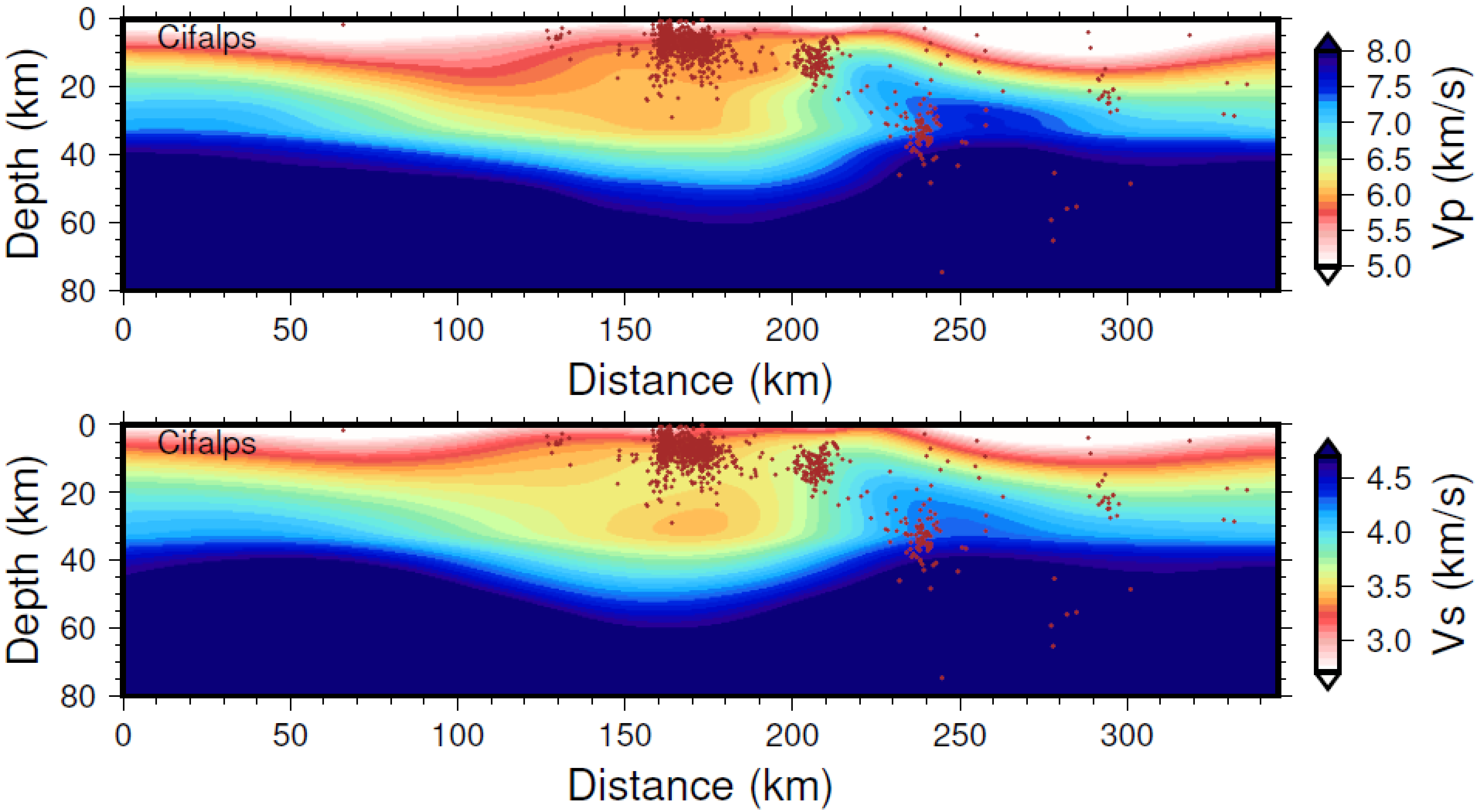
Upper layer sampling?



Initial location in red; Final location in blue
Shallower hypocenters in 3D final model

CifAlps cross-section

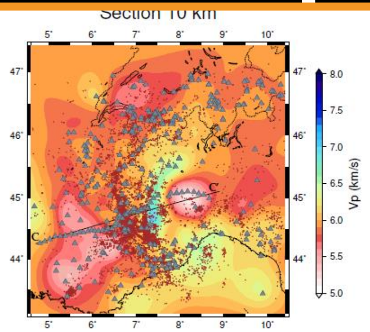




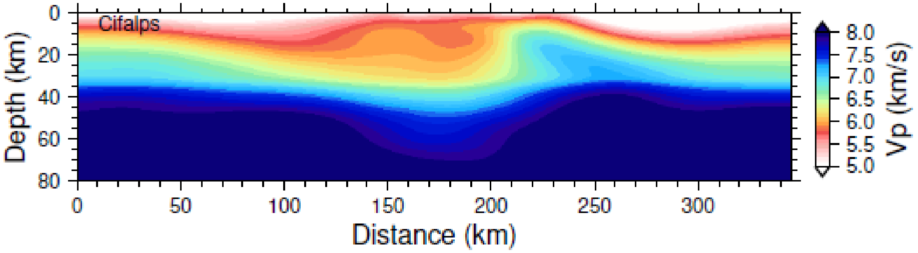
Known structural features such as the Ivrea body

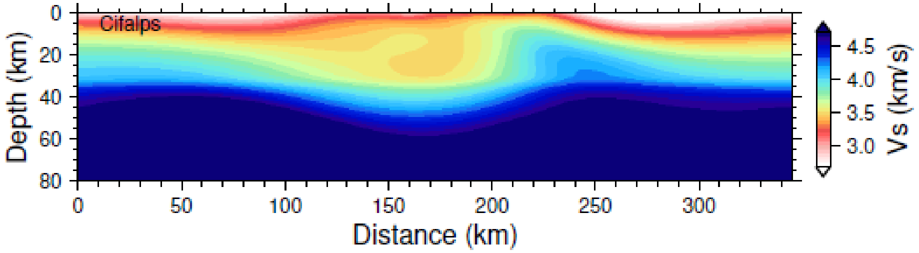
See paper
by Malusà et al (2021)
for deeper understanding

CifAlps cross-section: data & initial models

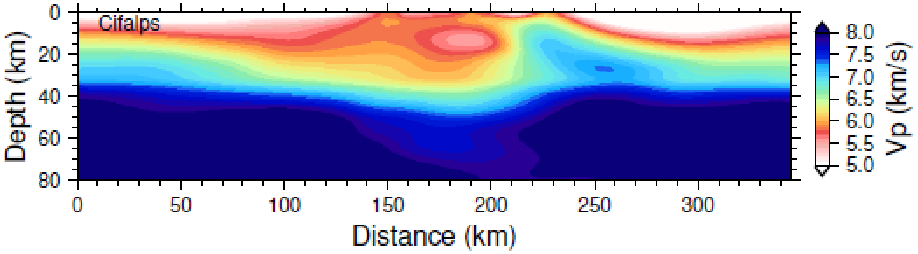


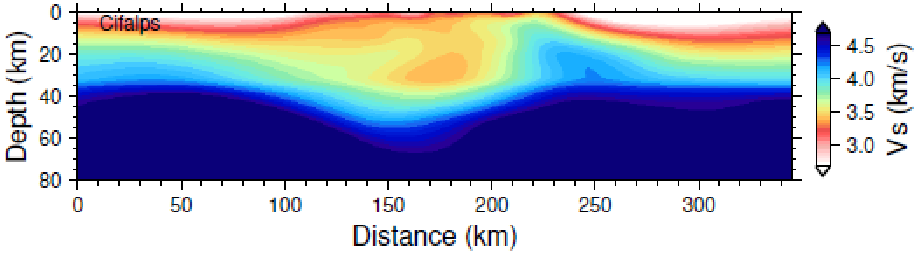
Full database
Standard scheme



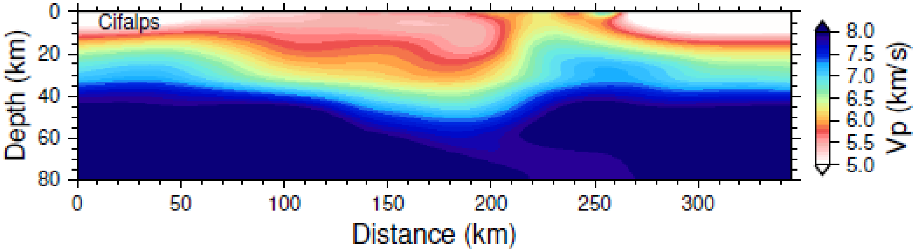


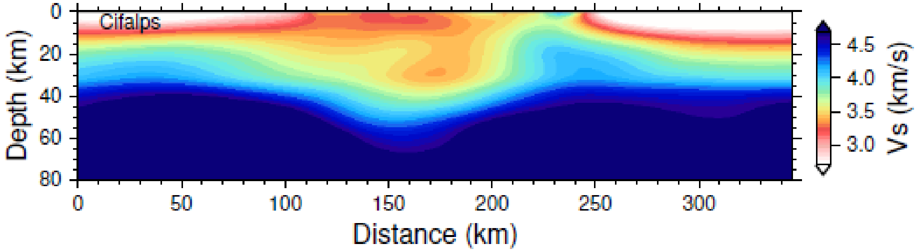
Selected Potin database
Standard scheme



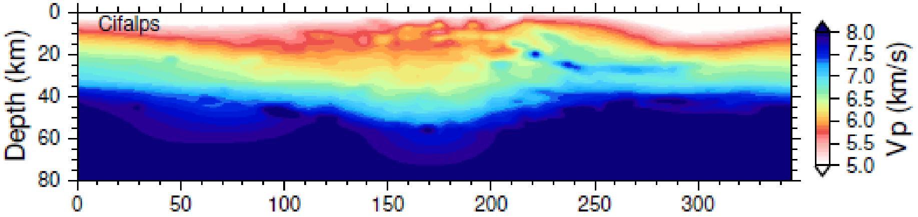


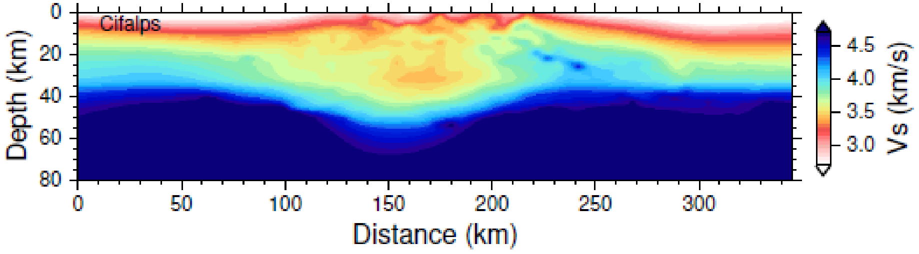
ANT-inspired initial model
Standard scheme





Only damped term
No smoothing term

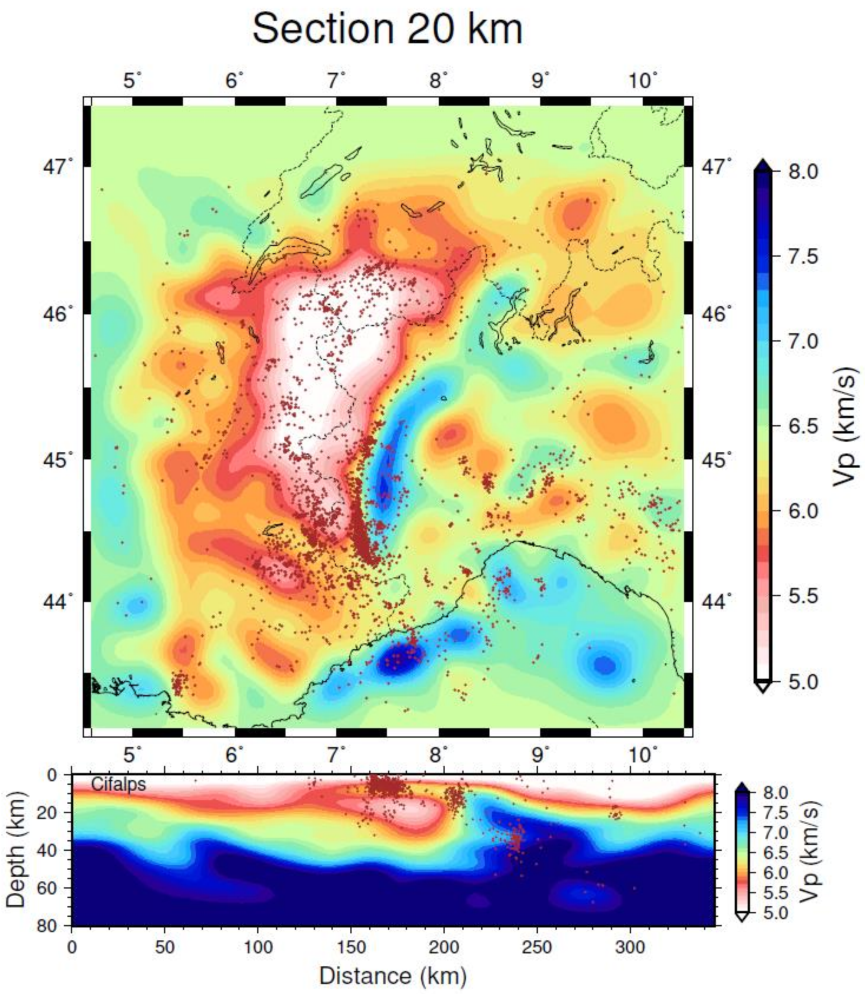




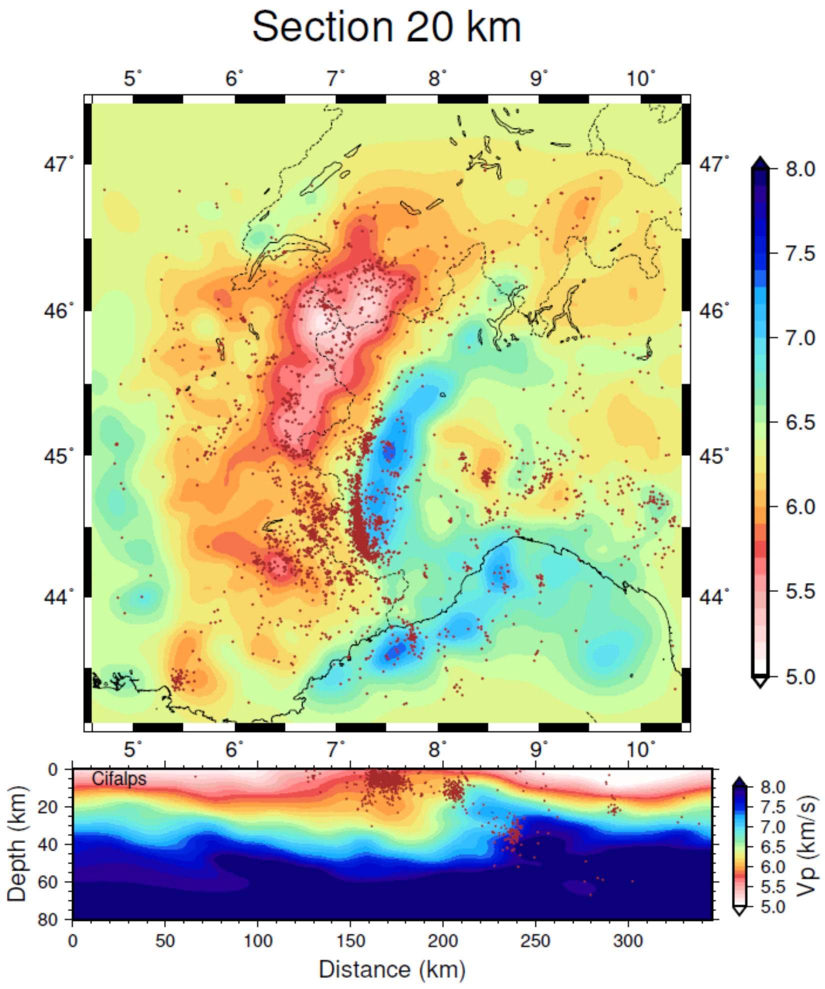
Never done in fact: > 1000 iterations

Distance (km)

Alternative model perturbation smoothing



Gaussian smoothing on model perturbation



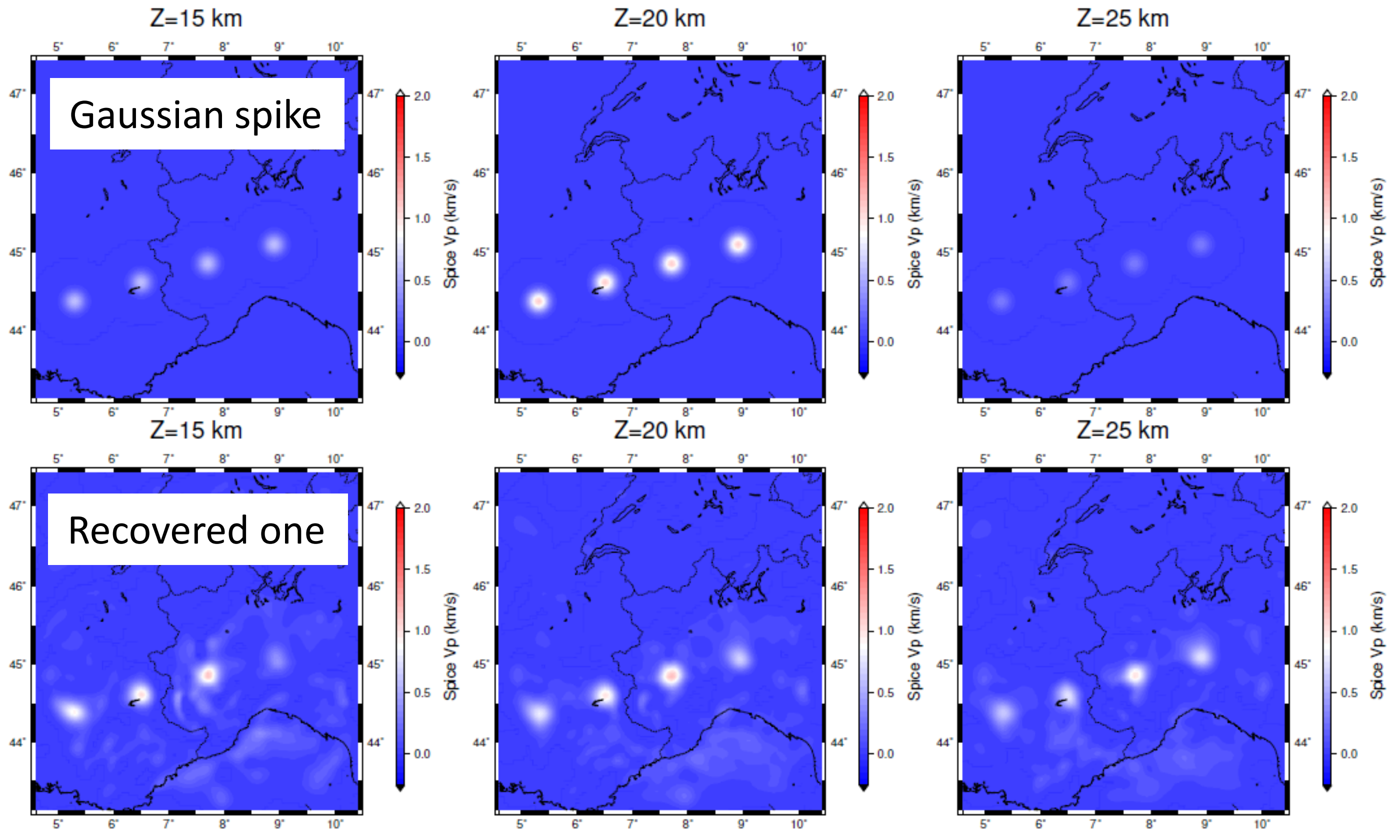
Total variation smoothing on model perturbation

Spike analysis: local resolution # model unicity

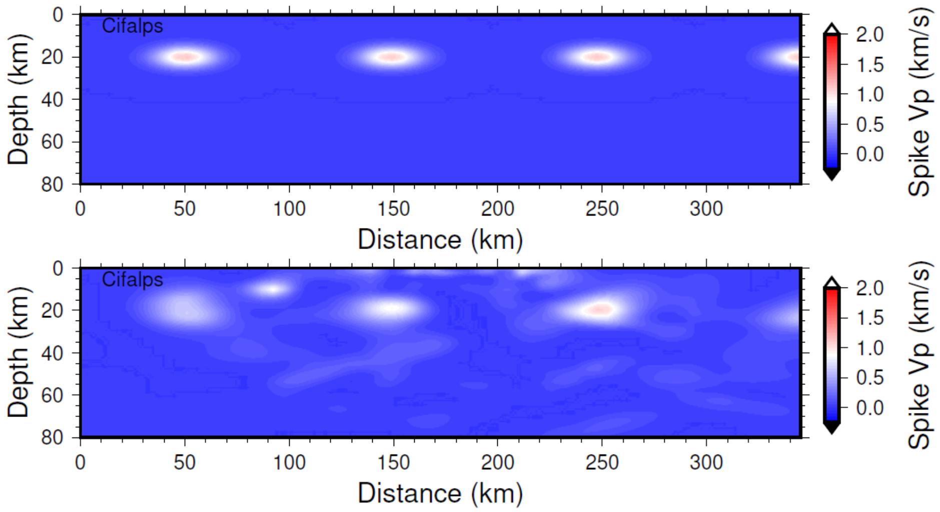
Different output models provide similar spike tests:
local resolution at the « optimal » model

Each spike reconstruction is independent: only the plotting combines them (#checkerboard test)

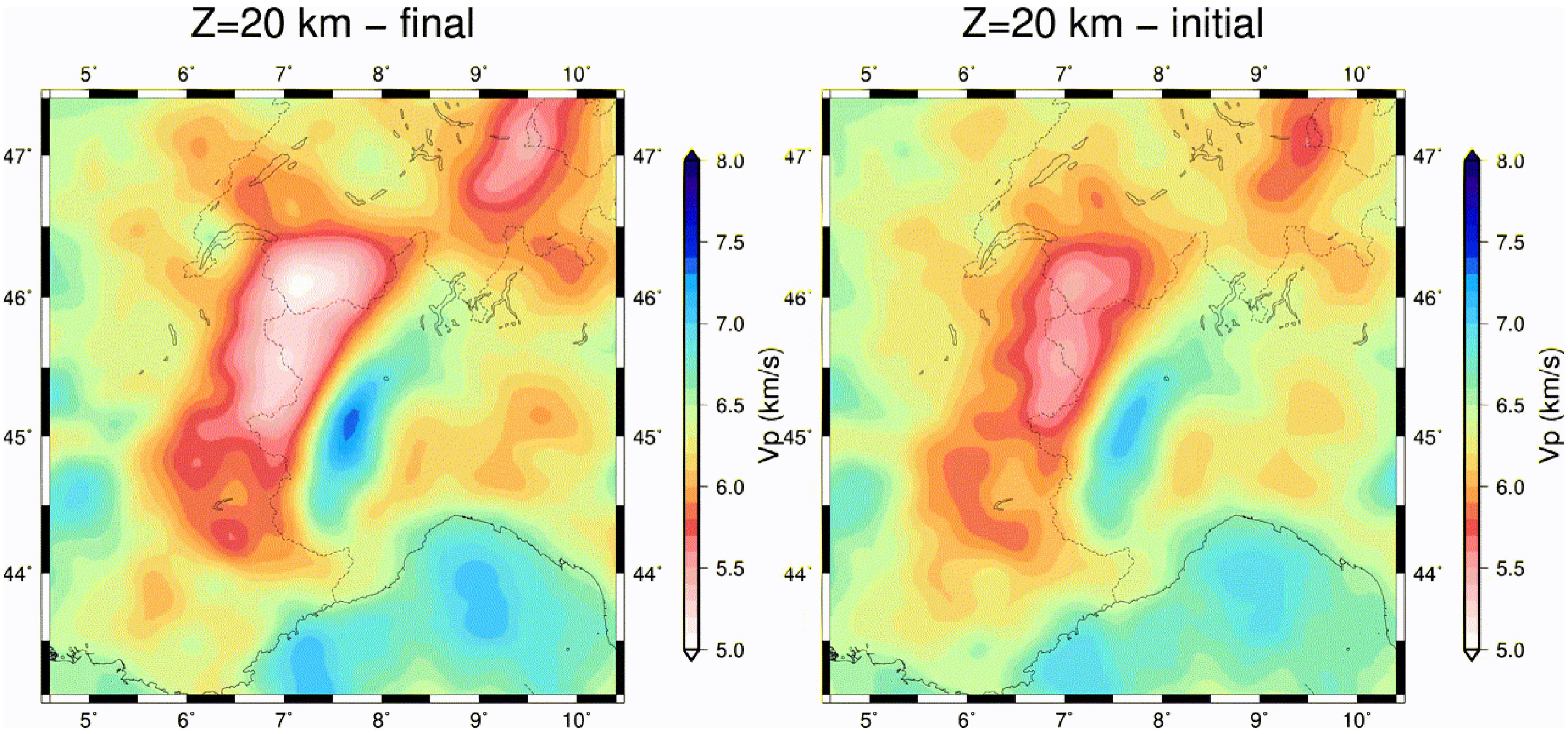
Horizontal section



Cifalps section



Take-away message of ray approach



Traveltime tomography is agnostic to frequency content of seismic waves.

Avoid over-interpretation of images
Support your interpretation with other reconstructions

Outline on first-arrival traveltime tomography

- Images at very different scales
- Waves and Phases: various concepts
- Few points on first-break ray-based tomography
- Illustration on 30-years Western Alps tomography
- First-break eikonal-based tomography
- First-break wave-equation-based tomography
- Hypocenter-velocity joint inversion
- Conclusion

Delayed eikonal-based tomography
based on model gradient building

Differential geometry: no need to introduce Lagrangian multipliers but such introduction leads to error-free manipulation

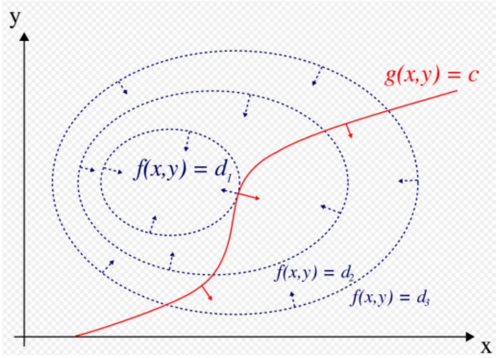
PDE-constrained optimization (alternative RD)

(see also Menke, 2012)

$$\min_s \frac{1}{N_{s,r}} \sum_{s,r} \frac{1}{2} (T_{obs} - \mathcal{R}(T(s)))^t \, (T_{obs} - \mathcal{R}(T(s)))$$

subject to $|\nabla T(x,y,z)| = s(x,y,z)$ or $\mathcal{H}(x,s,\nabla T_e)$=0

$s(x,y,z)$ is the slowness



Lagrangian multipliers approach (https://en.wikipedia.org/wiki/Lagrange_multiplier)

For the case of only one constraint and only two choice variables (as exemplified in Figure 1), consider the optimization problem

maximize $f(x, y)$
subject to $g(x, y) = 0$.

(Sometimes an additive constant is shown separately rather than being included in g, in which case the constraint is written $g(x, y) = c$, as in Figure 1.) We assume that both f and g have continuous first partial derivatives. We introduce a new variable (λ) called a **Lagrange multiplier** and study the **Lagrange function** (or **Lagrangian** or **Lagrangian expression**) defined by

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda \cdot g(x, y),$$

where the λ term may be either added or subtracted. If $f(x_0, y_0)$ is a maximum of $f(x, y)$ for the original constrained problem, then there exists λ_0 such that (x_0, y_0, λ_0) is a stationary point for the Lagrange function (stationary points are those points where the partial derivatives of $\mathcal{L}$ are zero). However, not all stationary points yield a solution of the original problem. Thus, the method of Lagrange multipliers yields a necessary condition for optimality in constrained problems.[3][4][5][6][7] Sufficient conditions for a minimum or maximum also exist, but if a particular candidate solution satisfies the sufficient conditions, it is only guaranteed that that solution is the best one *locally* – that is, it is better than any permissible nearby points. The *global* optimum can be found by comparing the values of the original objective function at the points satisfying the necessary and locally sufficient conditions.

The method of Lagrange multipliers relies on the intuition that at a maximum, $f(x, y)$ cannot be increasing in the direction of any neighboring point where $g = 0$. If it were, we could walk along $g = 0$ to get higher, meaning that the starting point wasn't actually the maximum.

wikipedia

$$\mathcal{L}(s,T,\lambda) = \underbrace{\frac{1}{N_{e,r}} \sum_{e,r} \frac{1}{2} (T_{obs} - \mathcal{R}[T_e])^t \, (T_{obs} - \mathcal{R}[T_e])} - \frac{1}{2} \sum_e \int_\Omega \lambda_e \, \mathcal{H}(x,s,\nabla T_e) dx$$

converted in a minimisation problem

Lagrangian formulation

$$\mathcal{L}(s,T,\lambda) = \underbrace{\frac{1}{N_{e,r}}\sum_{e,r}\frac{1}{2}(T_{obs}-\mathcal{R}[T_e])^t W_d^t W_d (T_{obs}-\mathcal{R}[T_e])}_{\text{Data misfit}} - \frac{1}{2}\sum_e \int_\Omega \lambda_e\, \mathcal{H}(x,s,\nabla T_e)dx$$

Equation-based constraint

where the operator $\mathcal{R}$ extracts synthetic times at receiver positions, fields $T(x)$, $\lambda(x)$ for one event and model parameter $s(x)$ are independent quantities (they are connected at the minimisation point, called realization point). Data are weighted by a matrix W_d

$\mathcal{L}(s,T,\lambda) = \mathcal{L}_1(s,T) + \mathcal{L}_2(s,T,\lambda)$: $\mathcal{L}_1$ has the same value as the misfit function $\mathcal{C}(s)$

The optimal solution will be given by $\mathrm{d}\mathcal{L} = 0$

$$\begin{cases} \dfrac{\partial \mathcal{L}^t}{\partial s}.\delta s = 0 \\ \dfrac{\partial \mathcal{L}^t}{\partial T}.\delta T = 0 \\ \dfrac{\partial \mathcal{L}^t}{\partial \lambda}.\delta \lambda = 0 \end{cases}$$

necessary conditions: KKT conditions (Karush-Kuhn-Tucker)

first-order optimality conditions in a full space search (s,T,λ) ☹

Reduced-model approach

If $\frac{\partial \mathcal{L}^t}{\partial \lambda}.\delta\lambda = 0 \rightarrow \mathcal{H}(\nabla \mathrm{T}) = 0 \rightarrow T^*$ ← Eikonal

and $\frac{\partial \mathcal{L}^t}{\partial T}.\delta T = 0 \rightarrow \frac{\partial \mathcal{L}_1^t}{\partial T}.\delta T + \frac{\partial \mathcal{L}_2^t}{\partial T}.\delta T = 0 \rightarrow \lambda^*,$ ← Adjoint

then $d\mathcal{L} = \frac{\partial \mathcal{L}^t\,(s,T^*(s),\lambda^*(s))}{\partial s}.\delta s = \frac{\partial \mathcal{C}}{\partial s}(s).\delta s$ ← Gradient

Reduced-model approach

If $\frac{\partial \mathcal{L}^t}{\partial \lambda}.\delta\lambda = 0 \rightarrow \mathcal{H}(\nabla \mathrm{T}) = 0 \rightarrow T^*$ — Eikonal

and $\frac{\partial \mathcal{L}^t}{\partial T}.\delta T = 0 \rightarrow \frac{\partial \mathcal{L}_1^t}{\partial T}.\delta T + \frac{\partial \mathcal{L}_2^t}{\partial T}.\delta T = 0 \rightarrow \lambda^*,$ — Adjoint

then $d\mathcal{L} = \frac{\partial \mathcal{L}^t\,(s,T^*(s),\lambda^*(s))}{\partial s}.\delta s = \frac{\partial \mathcal{C}}{\partial s}(s).\delta s$ — Gradient

For one source: $\mathcal{H}\big(\nabla(\mathrm{T} + \delta T)\big) = \mathcal{H}(\nabla T) + \frac{\partial \mathcal{H}}{\partial \nabla T}.\nabla \delta T$

$$\left(\frac{\partial \mathcal{H}}{\partial \nabla_x T}, \frac{\partial \mathcal{H}}{\partial \nabla_y T}, \frac{\partial \mathcal{H}}{\partial \nabla_z T}\right)$$

$\frac{\partial \mathcal{L}_2^t}{\partial T}.\delta T$?

$$\frac{\partial \mathcal{L}_2^t}{\partial T}.\delta T = -\frac{1}{2}\int_\Omega \lambda \frac{\partial \mathcal{H}}{\partial \nabla T}.\nabla \delta T dx$$

Divergence theorem: $\nabla.(\delta T \lambda \frac{\partial \mathcal{H}}{\partial \nabla T}) = \delta T\, \nabla.\left(\lambda \frac{\partial \mathcal{H}}{\partial \nabla T}\right) + \nabla \delta T.\lambda \frac{\partial \mathcal{H}}{\partial \nabla T}$

$$\frac{\partial \mathcal{L}_2^t}{\partial T}.\delta T = -\frac{1}{2}\int_\Omega \nabla.(\delta T \lambda \frac{\partial \mathcal{H}}{\partial \nabla T})dx + \frac{1}{2}\int_\Omega \delta T \nabla.(\lambda \frac{\partial \mathcal{H}}{\partial \nabla T})dx$$

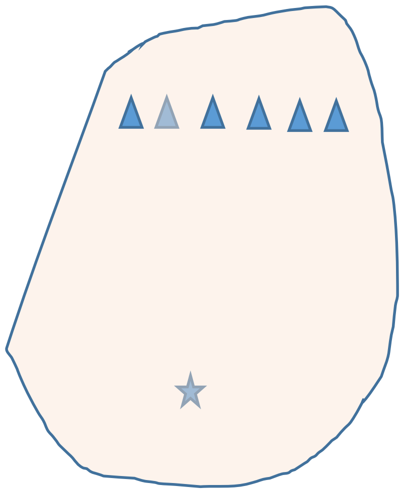
$$\frac{\partial \mathcal{L}_2^t}{\partial T}.\delta T = -\frac{1}{2}\int_{\partial\Omega} \delta T \lambda \frac{\partial \mathcal{H}}{\partial \nabla T} dl + \frac{1}{2}\int_\Omega \delta T \nabla.(\lambda \frac{\partial \mathcal{H}}{\partial \nabla T})dx$$

Differential manipulation

Adjoint PDE: transport equation

$$\frac{\partial \mathcal{L}^t}{\partial T}.\delta T = \frac{1}{N_{e,r}}\sum_{e,r}(T_{obs}-\mathcal{R}[T_e(s)])^t W_d^t\,.\,\delta T - \frac{1}{2}\sum_e \int_{\partial\Omega} \cancel{\delta T\lambda_e \frac{\partial\mathcal{H}}{\partial\nabla T}\, dl} + \frac{1}{2}\int_\Omega \delta T \nabla.\left(\lambda_e \frac{\partial \mathcal{H}}{\partial \nabla T}\right)dx = 0$$

No restriction if Dirichlet condition ($\lambda = 0$ over $\partial\Omega$)



For one event: $\int_\Omega (T_{obs}-\mathcal{R}[T(s)])^t W_d^t\,\delta(x-x_r)\delta T dx$

$$\int_\Omega \delta T\left[\nabla.\left(\lambda\frac{\partial\mathcal{H}}{\partial\nabla T}\right) + (T_{obs}-\mathcal{R}[T(s)])^t W_d^t\,\delta(x-x_r)\right]dx = 0 \quad \forall\, \delta T$$

Transport equation (similar to the one for the amplitude)

Adjoint PDE: transport equation

$$\mathcal{L}(s,T,\lambda) = \underbrace{\frac{1}{N_{e,r}} \sum_{e,r} \frac{1}{2} (T_{obs} - \mathcal{R}[T_e])^t W_d^t W_d (T_{obs} - \mathcal{R}[T_e])} - \frac{1}{2} \sum_e \int_\Omega \lambda_e \, \mathcal{H}(x,s,\nabla T_e) dx$$

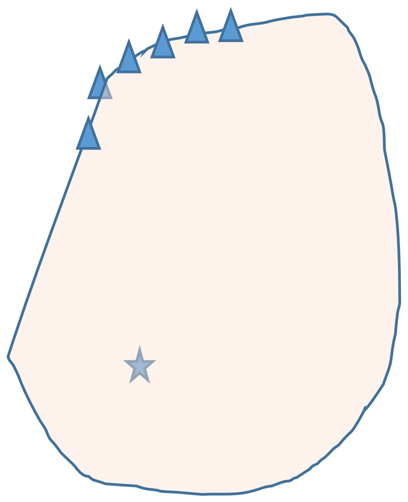
$$\left[\nabla . \left(\lambda(x) \frac{\partial \mathcal{H}}{\partial \nabla T} \right) = - \sum_r (T_{obs} - \mathcal{R}[T(s)])^t W_d^t \, \delta(x - x_r) \right]$$

Reduced-parameter $\lambda(x)$

Linear PDE for adjoint quantity $\lambda(x)$ for one event but all associated receivers

Adjoint/Lagrangian approach: a simple way for differential geometry …

Adjoint PDE: boundary conditions



(Leung & Qian, 2006; Taillandier et al, 2009; Waheed et al, 2016)

$$\nabla.\left(\lambda(x)\frac{\partial\mathcal{H}}{\partial\nabla T}\right)=0$$

$$\lambda(x)\frac{\partial\mathcal{H}}{\partial\nabla T}\bigg|_{\partial\Omega}=\sum_e\left(T_{obs}-\mathcal{R}[T(s)]\right)^t C_d^t$$

Continuity of boundary conditions?

RHS injection (even if singular – point dirac injection) is easier than boundary conditions (missing values)

(limited aperture of receiver distribution)

especially when considering other observables (Tavakoli F. et al, 2017)

Model gradient estimation

independent variables s, T, λ

$$\mathcal{L}(s,T,\lambda) = \frac{1}{N_{e,r}} \sum_{e,r} \frac{1}{2} (T_{obs} - \mathcal{R}[T])^t W_d^T W_d (T_{obs} - \mathcal{R}[T]) - \frac{1}{2} \sum_e \int_\Omega \lambda\, \mathcal{H}(x, \nabla T, s) dx$$

- ❑ Isotropic model: $\mathcal{H}(x, \nabla T, s) = |\nabla T|^2 - s(x)^2$ for one event

$$\frac{\partial \mathcal{C}}{\partial s}(s) = \left.\frac{\partial \mathcal{L}(s,T,\lambda)}{\partial s}\right|_{s,T^*,\lambda^*} = \sum_e \int_\Omega \lambda^*(x) s(x) dx$$

T^* *solution of the Eikonal equation*

λ^* *solution of the transport equation* (connected to T^*)

Model slowness $s(x)$?

Model gradient estimation (anisotropy)

$$\mathcal{L}(s,T,\lambda) = \frac{1}{N_{e,r}} \sum_{e,r} \frac{1}{2} (T_{obs} - \mathcal{R}[T])^t W_d^T W_d (T_{obs} - \mathcal{R}[T]) - \frac{1}{2} \sum_e \int_\Omega \lambda\, \mathcal{H}(x, \nabla T, s) dx$$

- Anisotropic model: $m(v_v, \varepsilon, \delta)$ $- A(x,m), B(x,m), E(x,m)$ *closed-form expression*

$$\rightarrow \mathcal{H}(x,\nabla T) = A(x, v_v, \varepsilon, \delta)(\nabla_x T)^2 + B(x, v_v, \varepsilon, \delta)(\nabla_z T)^2 + E(x, v_v, \varepsilon, \delta)(\nabla_x T)^2 (\nabla_z T)^2 - 1 = 0$$

(Alkhalifah, 2003; Waheed, 2014; Tavakoli F., 2017; Le Bouteiller, 2018)

$$\frac{\partial \mathcal{C}}{\partial m}(v_v, \varepsilon, \delta\,) = \left.\frac{\partial \mathcal{L}(v_v, \varepsilon, \delta\,, T, \lambda)}{\partial m}\right|_{s,T^*,\lambda^*} = \sum_e \int_\Omega \lambda^*(x) \frac{\partial \mathcal{H}(\nabla T^*(s))}{\partial m} dx$$

$$\frac{\partial \mathcal{H}(\nabla T(s))}{\partial v_v} ?$$

analytical expression easy to compute from $A(x,m), B(x,m), E(x,m)$

Adjoint PDE from reverse differentiation

$$\left[\nabla.\left(\lambda(x)\,\frac{\partial\mathcal{H}}{\partial\nabla T}\right) = -\sum_r \left(T_{obs} - \mathcal{R}[T(s)]\right)^t W_d^t\,\delta(x - x_r)\right]$$

Adjoint/Lagrangian approach: a simple way for differential geometry …

Connection with the so-called reverse differentiation
related to the chain rule for derivatives

Reverse differentiation fashionable concept especially for AI applications
with automatic differentiation approaches

(see illustrations on internet by J.-M. Mirebeau with notebooks)

Reverse differentiation

What is reverse differentiation?

$slowness\ u = {}^{1}/_{c}$

Forward differentiation (more familiar concept): FD

$$T(x, u + \varepsilon\delta u) = T(x, u) + \varepsilon\, \mu(x, u)$$

linear approximation μ (tangent information) for the time perturbation

Reverse differentiation (maybe less familiar concept): RD

$$T(x, u + \varepsilon\delta u) = T(x, u) + \varepsilon \int_{\Omega} \varrho(x, u)\delta u$$

Linear approximation ϱ (tangent information) for the slowness perturbation over Ω

Numerical **consistent** relation between time perturbation and slowness perturbation! $\mu(x, u) = \int_{\Omega} \varrho(x, u)\delta u$

FD & RD perturbation versus ray perturbation

FD: **slowness perturbation -> time perturbation**

For a given point x, $\delta u(x) \rightarrow \delta T(.)$ everywhere

RD: **time perturbation -> slowness perturbation**

For a given point x, $\delta T(x) \rightarrow \delta u(.)$ everywhere

i.e. time residual at a given station

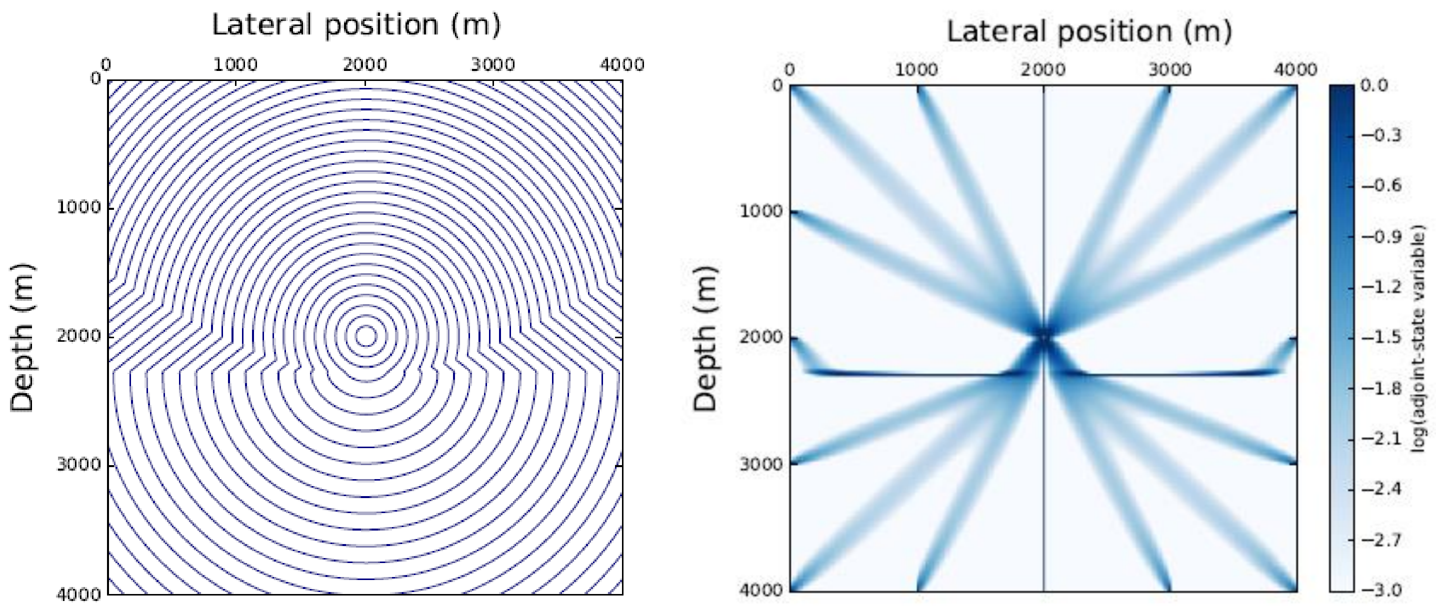
Ray linear relation

$$T(x, u + \varepsilon \delta u) = T(x, u) + \varepsilon \int_{source}^{receiver} \delta u \, dl$$

Locally, the ray sensitivity is the length of the ray inside the cell of the grid

Eikonal sensitivity kernel (SK)

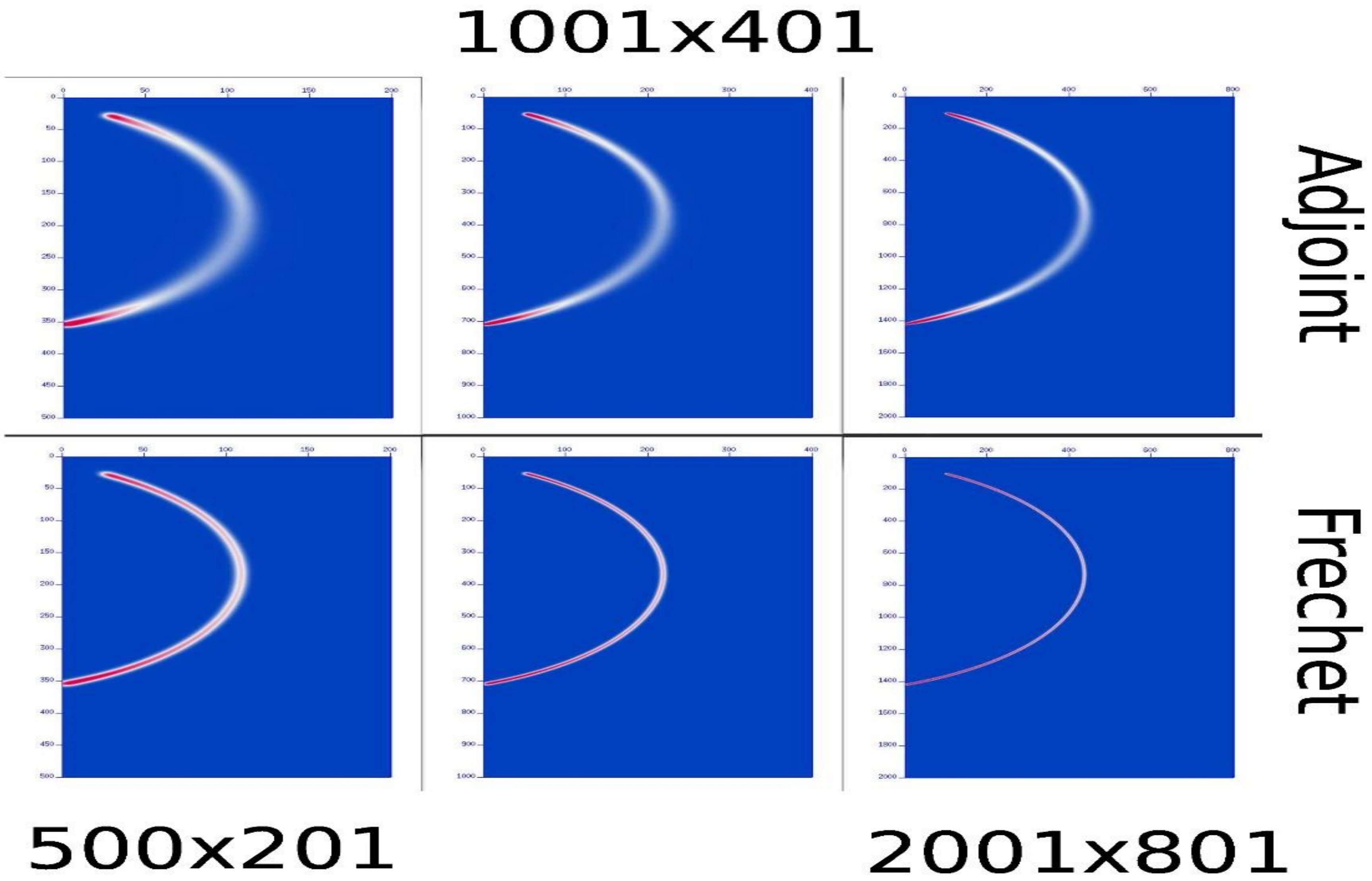
(Le Bouteiller, 2019)



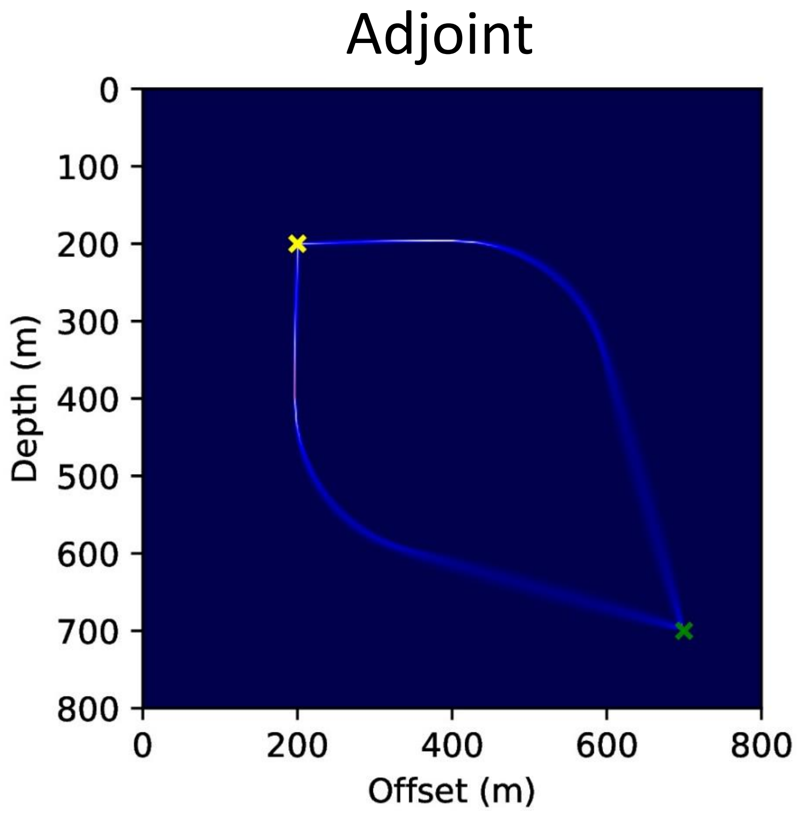
RD:
SK field

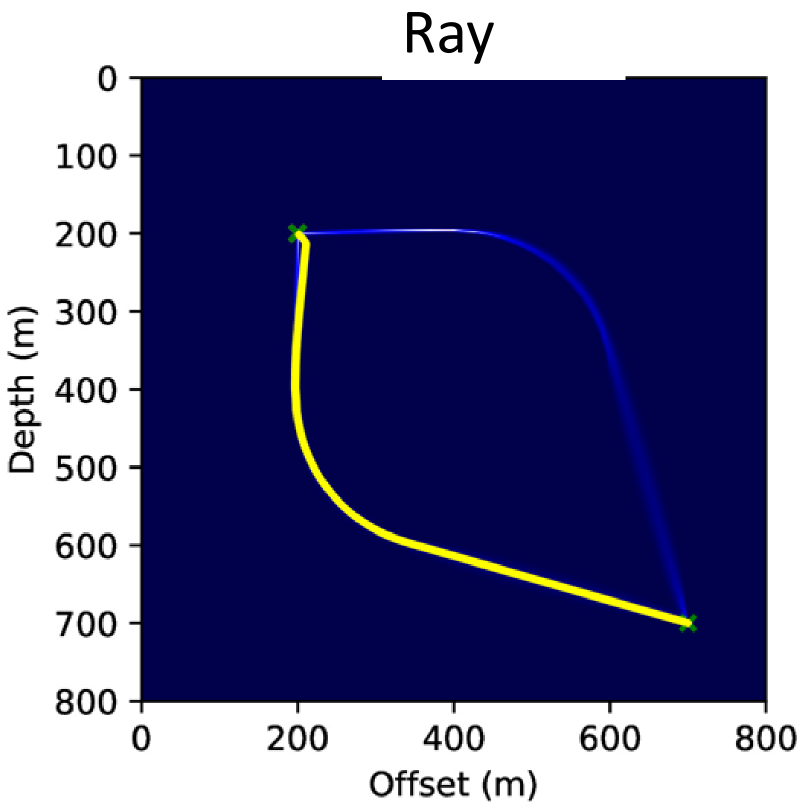
Reverse differentiation gives Eikonal sensitivity kernels:
where to insert velocity anomalies to match time data (sum of kernels over receivers) (Taillandier, 2009; Lelièvre et al, 2011; Tavakoli F. et al, 2017, 2019; Sambolian et al, 2019, 2021; Tong, 2021)
Sensitivity kernel defines zones of velocity perturbation affecting the time/phase at the receiver (agnostic to the frequency content of waves!)

RD linear PDEs



Cut-locus points





Cut-locus points: points where two rays provide same time:
Mathematical curiosity helping to understand the spreading of adjoint solution
(illustration of Fermat principle misleading interpretation)

Explicit or automatic differentiation

Differential geometry: no need to introduce Lagrangian multipliers but such introduction leads to error-free manipulation

However, automatic differentiation based on the *(reverse)* chain rule might ease the manipulation!

FD & RD linear PDEs

FD: **raw implementation** (linear PDE)

$$\nabla\mu\,(x).\nabla T(x) = u(x)\,\delta u(x)$$

Linearized Eikonal equation (Aldrige, 1994; Fomel, 2001; Franklin & Harris, 2001; Alkhalifah, 2002; Alkhalifah & Fomel, 2010; Li & Fomel, 2013)

FD: automatic differentiation (AD) implementation

Reverse chain rule (intermediate results to be saved) (Mirebeau & Dreo, 2017)

RD: **raw implementation** (linear PDE)

$$\nabla.[\varrho\,(x)\nabla T(x)] = 0 \quad in\ \ \Omega$$
$$s.t.$$
$$\varrho\,(x)[n(x).\,\nabla T(x)] = \delta T \ \text{ on } \delta\Omega$$

Adjoint transport equation (Leung & Qian, 2006; Taillandier, 2009; Tavakoli F et al, 2015, Tong, 2021)

RD: automatic differentiation (AD) implementation

Reverse chain rule (intermediate results to be saved) (Mirebeau & Portegies, 2019)

On the use of FD & RD PDEs solvers

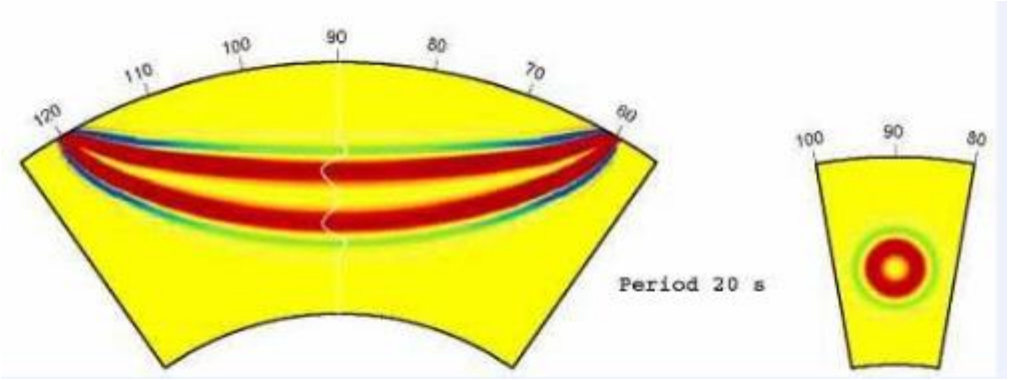
First-arrival traveltime tomography: we need

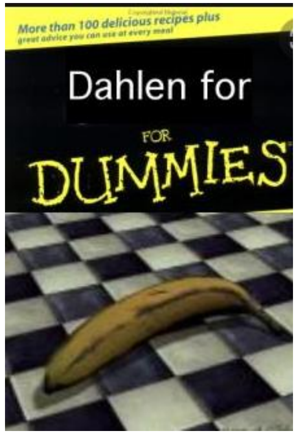
- Eikonal solver (time; non-linear PDE)
- RD solver (time SK; linear PDE) Péclet number is infinite

Computer complexity (non-linear, singularities; CPU, memory)…

Banana-Donut debate

https://www.geoazur.fr/GLOBALSEIS/nolet/BDdiscussion.html





(Dahlen et al, 2000; Dahlen & Nolet, 2005; Nolet, 2008)

Phase/time tomography based on viscous solution does not provide the so-called BD shape promoted by Dahlen & Nolet … !!!

Such shape is related to **ray concept** (and not to the **time/phase concept** for which the zero-sensitivity along the ray does not exist)

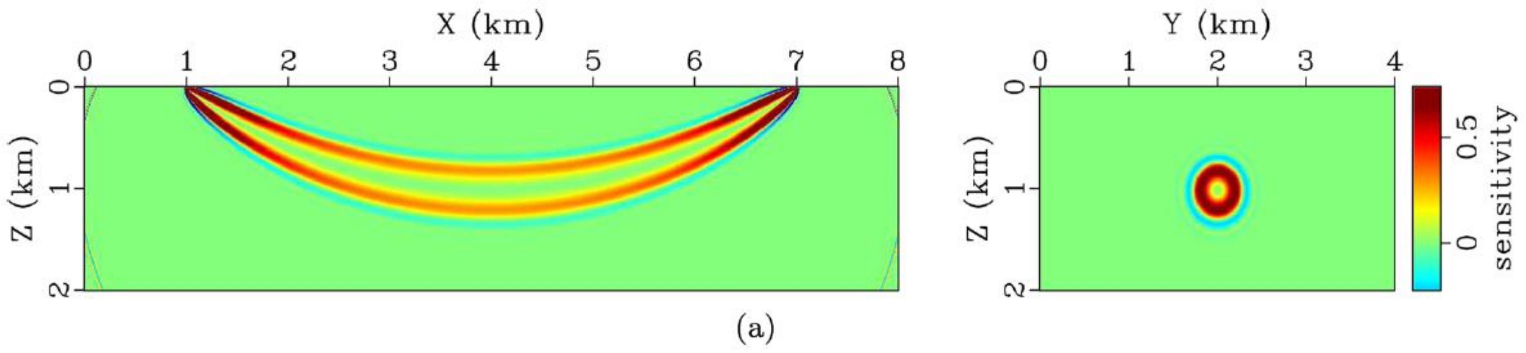
Many authors have tried to understand better this theoretical contradiction …

My understanding of this debate …

When considering phase information

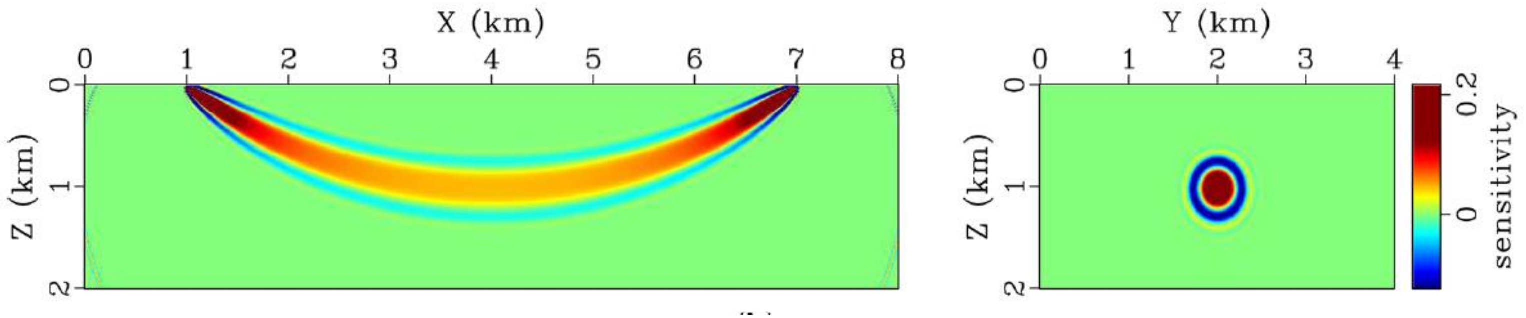
DIFFRACTION EFFECT NOT INCLUDED IN RAY APPROACH
DIFFRACTION EFFECT CAN BE INCLUDED IN THE SO-CALLED EIKONAL APPROACH

RAY SK



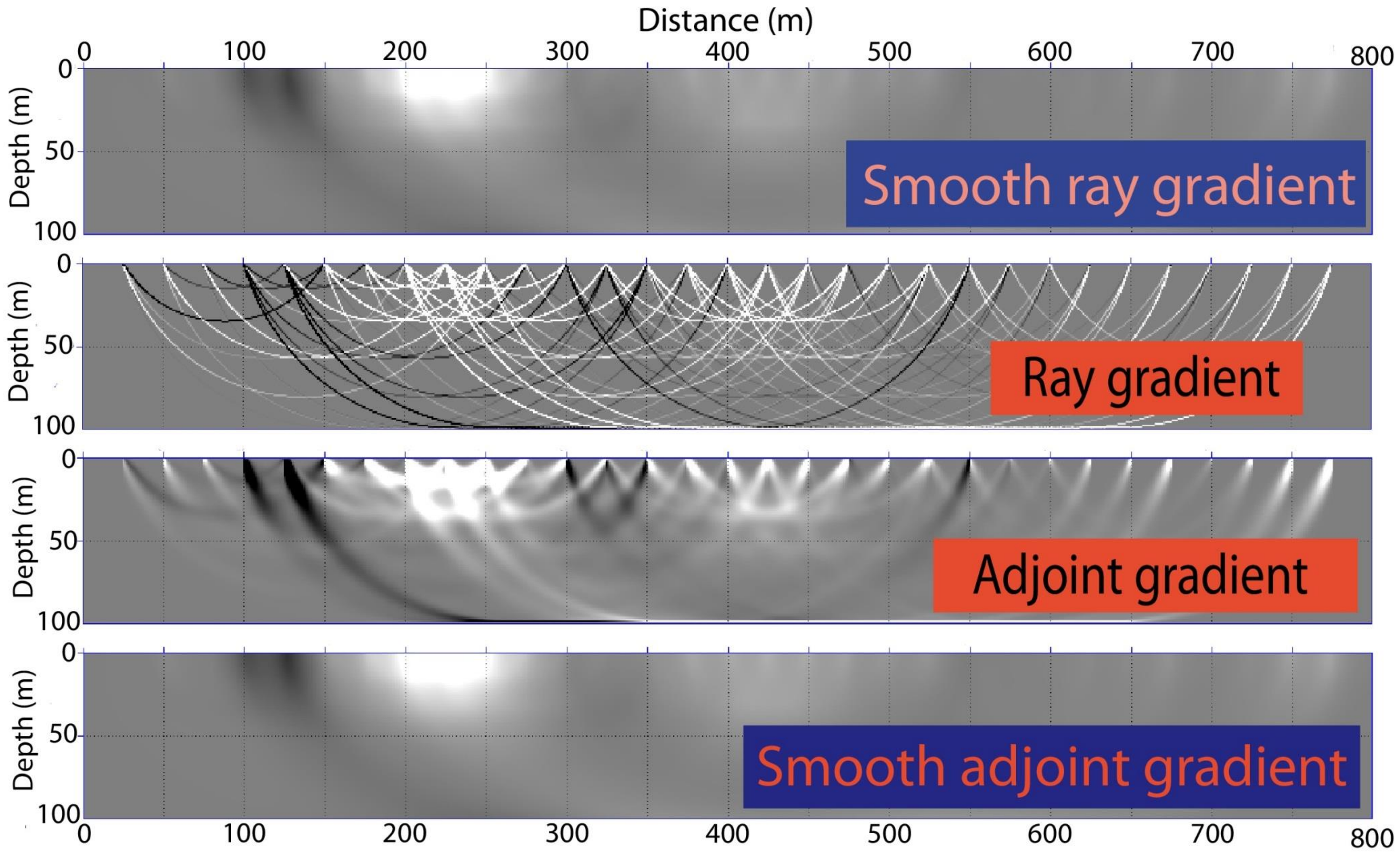
(From Djebbi & Alkhalifa, 2014)

EIK. SK



Outcome of this debate: is it really important for applications?

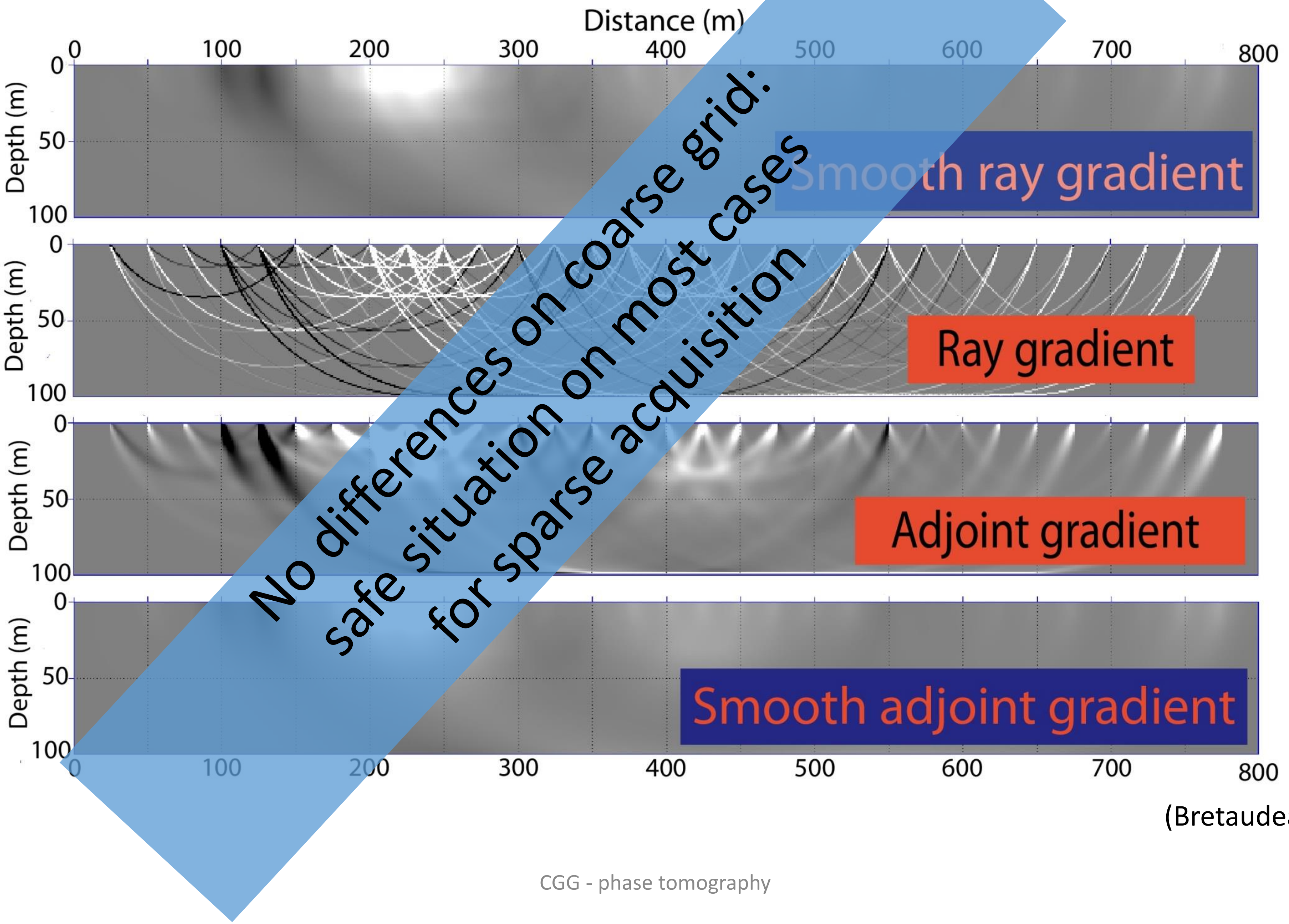
Projection of gradients on discrete model



Eikonal+Adjoint

(Bretaudeau et al, 2014)

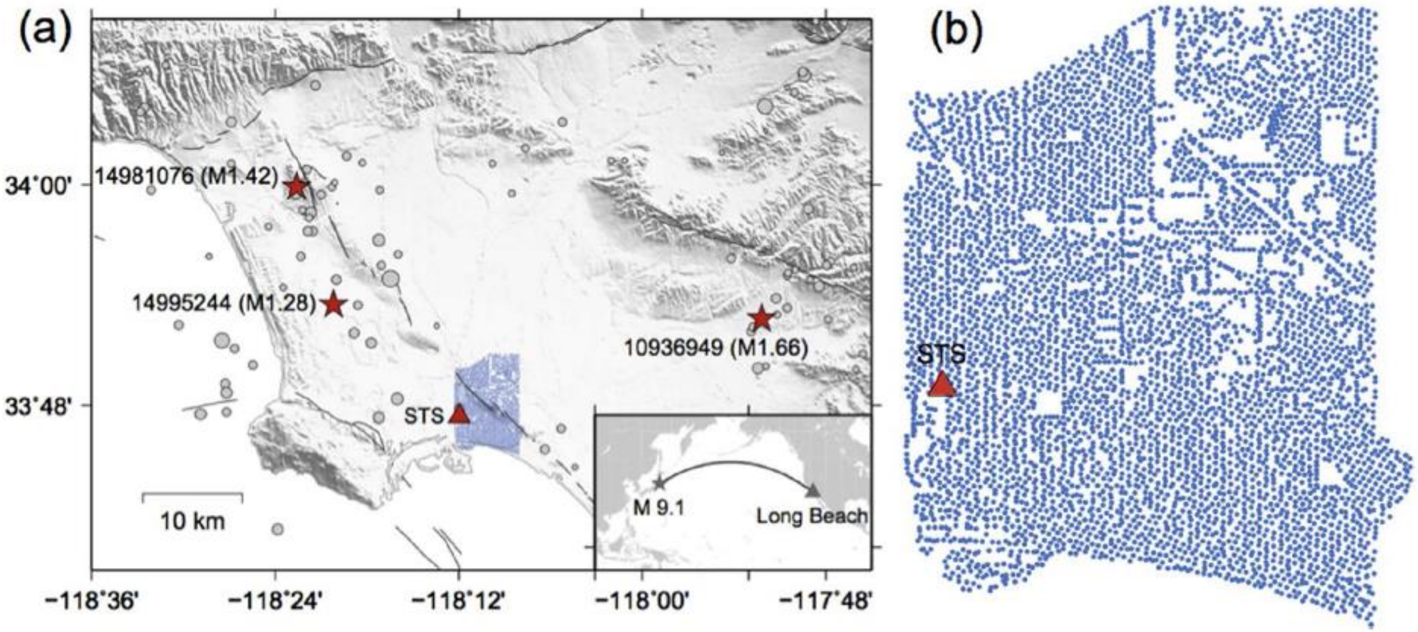
Projection of gradients on discrete model



(Bretaudeau et al, 2014)

Ray SK versus Eikonal SK?

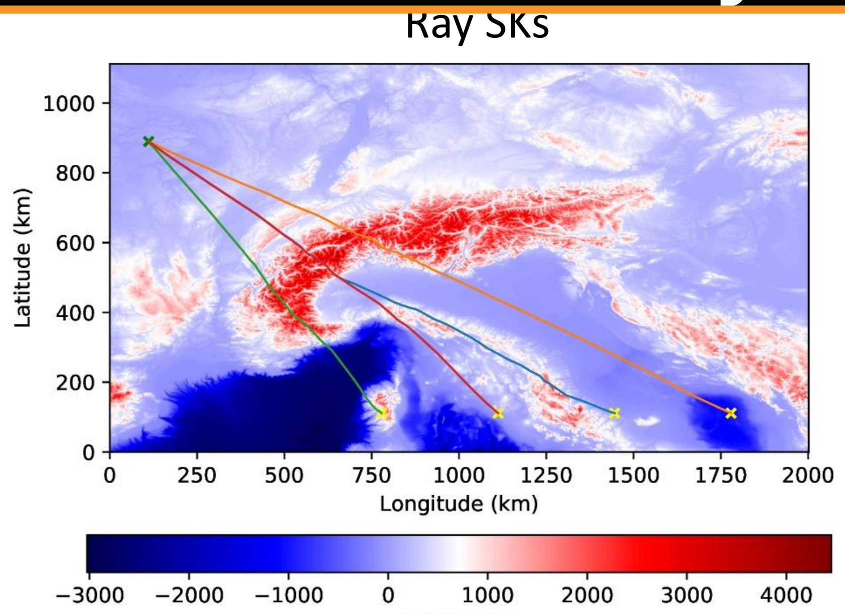
Does Eikonal-based tomography provide improved images than ray-based tomography?

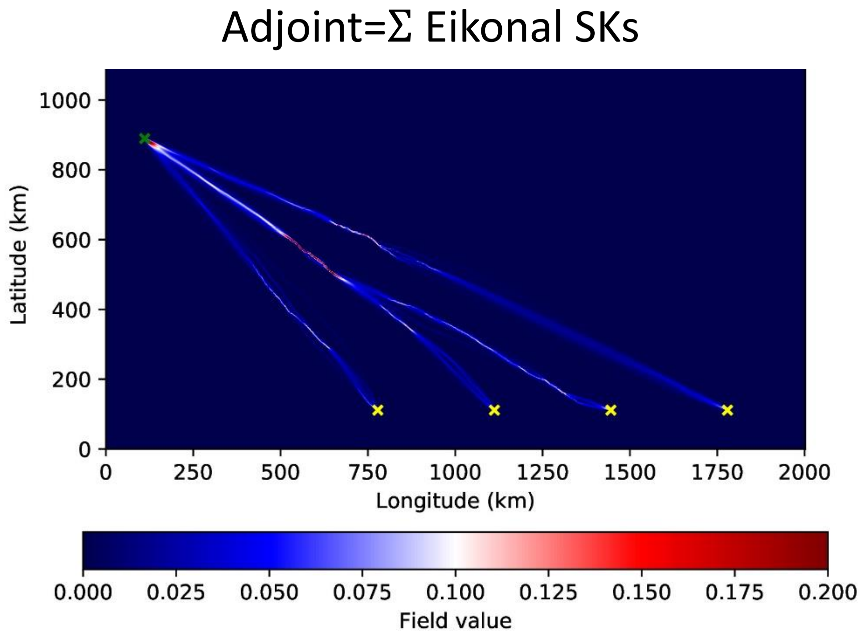


How to take benefit of such data density?

My negative answer (2014) has moved to a more positive answer (2019), especially when considering drastic increase of the acquisition density

Is it still true with array densification?





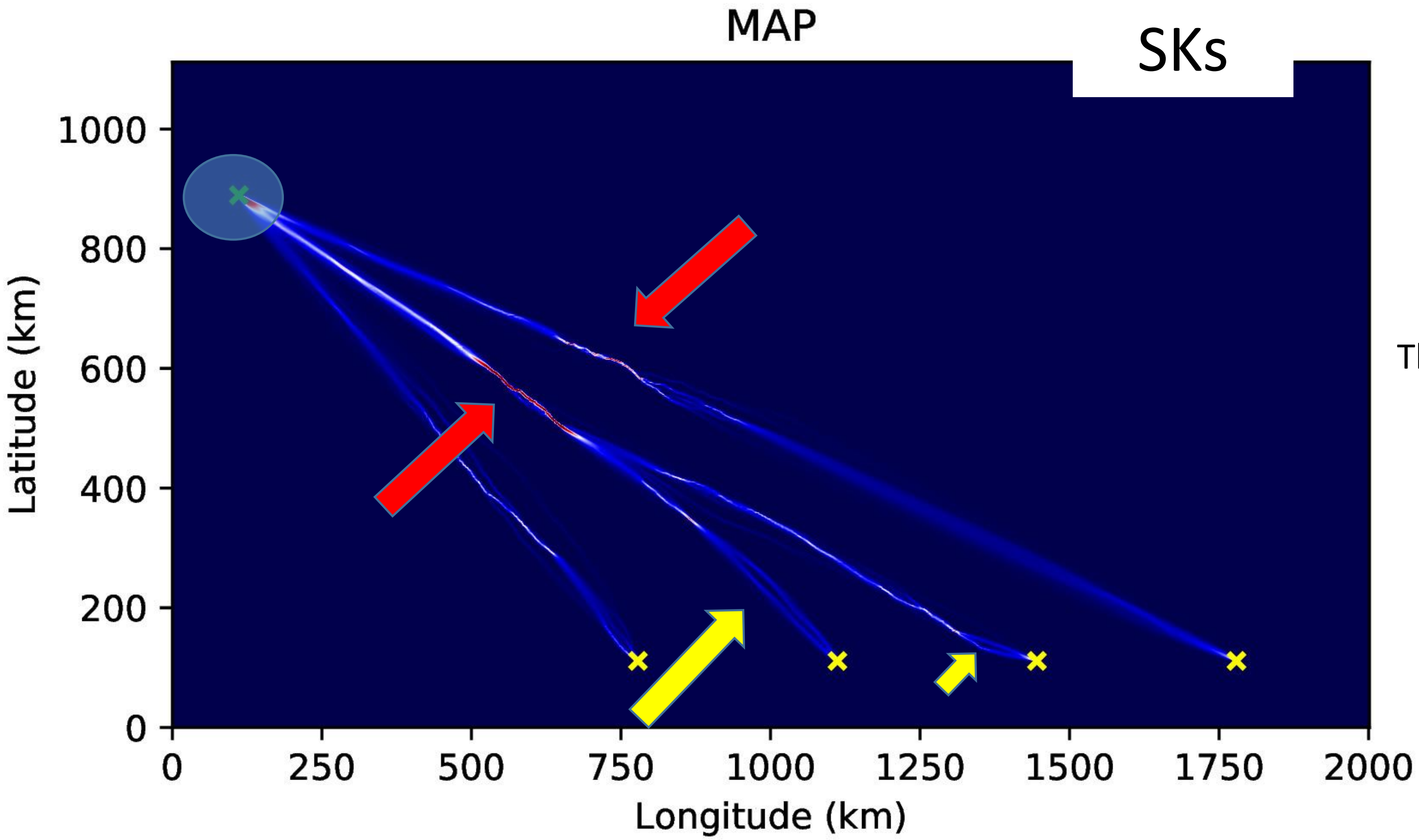
Toy example:
topography taken as velocity perturbation
(mimicking HF pattern)

Rays: same value (ray length) of the sensitivity
(velocity impacts only the ray path itself)

Eikonal SK exhibits more complex pattern: variable values along the trajectory !!!

SKs: different values with possible different paths before reaching the station

Eikonal SK: modulation & multi-paths



modulation

They do not exist in ray approach

multi-paths

Delayed Eikonal-based algorithm

Linearized Inverse Problem

One-loop procedure

Initial velocity model s_0

$k = 0$

One loop

$k = k + 1$

Eikonal solver $time(x)$ (src)

Transport solver $adjoint(x)$ (src)

Gradient estimation $\gamma_k = adjoint(x) * s(x)$

Model gradient smoothing $\tilde{\gamma}_k$

$\mathcal{H} ds_k = -\tilde{\gamma}_k$

$ds_k small$

End

Gradient methods/
Quasi-Newton methods such as l-BFGS method

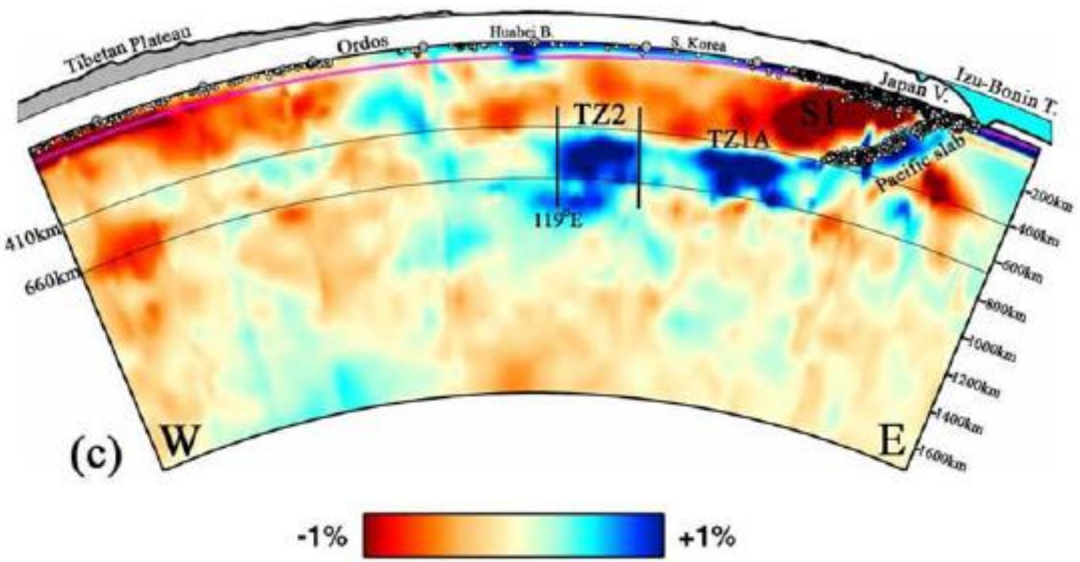
Storage of previous gradients and models

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

DRT versus DET: spatial difference

Delayed ray-based Tomography

$$\delta \mathrm{T}(\mathrm{s},\mathrm{r}) = \int \delta u(x(l))dl = \iiint \delta u(x)\delta\big(x - x(l)\big)K(s,r,x)dv$$



Still DRT provides impressive images while we do believe that DET would provide better images in the future, thanks to the densification of the available data.

(Li & van der Hilst, 2010)

Delayed Eikonal-based Tomography?

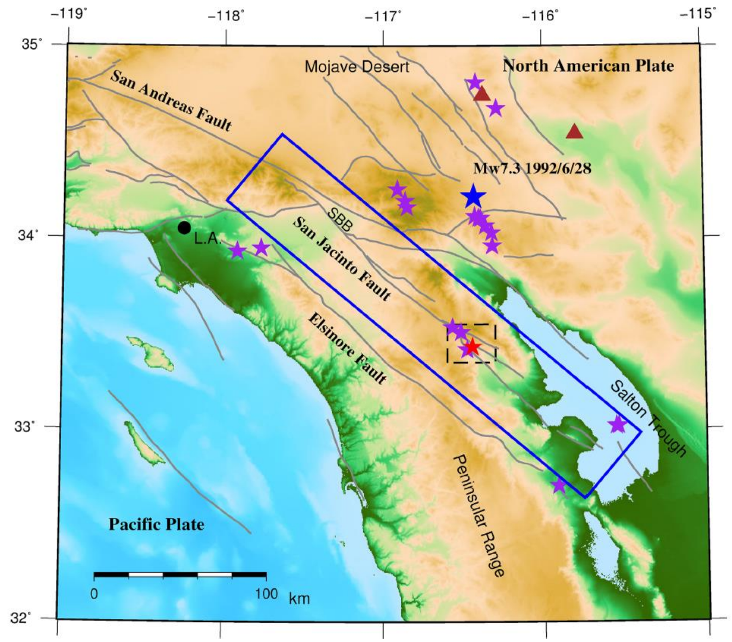
$$\delta \mathrm{T}(s,r) = \iiint \delta u(x)\, K(s,r,x)dv$$

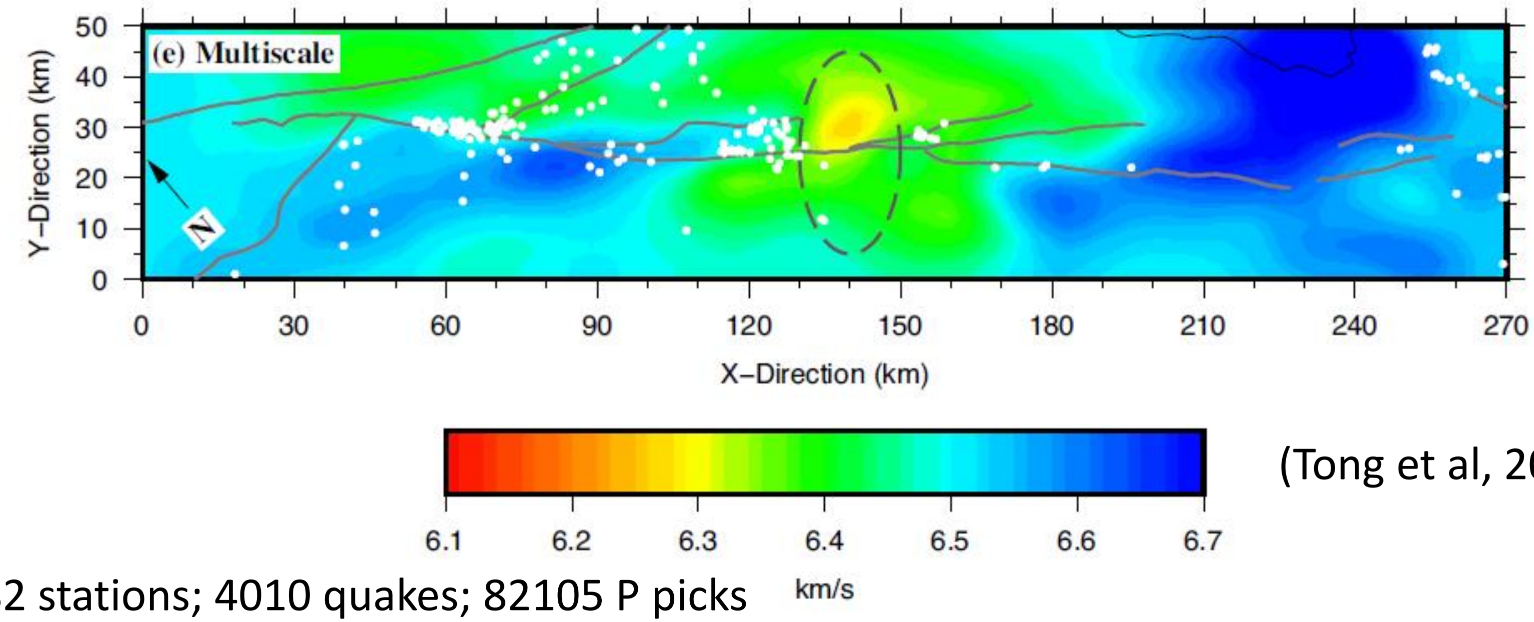
Volumic $K(s,r,x)$ still frequency-independent

Delayed Eikonal-based Tomography?

$$\delta \mathrm{T}(s,r) = \iiint \delta u(x)\, K(s,r,x) dv$$

Volumic $K(s,r,x)$ still frequency-independent





(Tong et al, 2019)

182 stations; 4010 quakes; 82105 P picks

DAS perspective: dense kinematic observables?

DRT versus DET: winner?

Delayed ray-based Tomography

$$\delta \mathrm{T}(\mathrm{s},\mathrm{r}) = \int \delta u(x(l))dl = \iiint \delta u(x)\delta\big(x - x(l)\big)K(s,r,x)dv$$

Delayed Eikonal-based Tomography?

$$\delta \mathrm{T}(s,r) = \iiint \delta u(x)\, K(s,r,x)dv$$

Volumic $K(s,r,x)$ still frequency-independent

Delayed Eikonal-based tomography has the same computational complexity than Delayed Ray-based tomography (asymptotic framework versus high-frequency asymptotic framework).

Both are agnostic to frequency content of seismic waves … **blue-sky information.**

These approaches are needed in order to start wave-equation tomography which is sensitive to frequency content, but which is subjected to cycle-skipping issues. Moreover, wave-equation tomography requires significant computer resources because the wave (elastodynamic) equation has to be solved.

Outline on first-arrival traveltime tomography

- Images at very different scales
- Waves and Phases: various concepts
- Few points on first-break ray-based tomography
- Illustration on 30-years Western Alps tomography
- First-break eikonal-based tomography
- First-break wave-equation-based tomography
- Hypocenter-velocity joint inversion
- Conclusion

Wave-equation tomography

Wave Equation Tomography (WET)

(Woodward, 1989; Luo & Schuster, 1991; Woodward, 1992, Van Leeuwen & Mulder, 2010)

Somehow immune to amplitude effects

- Time windowing strategy for waveform extraction (Maggi et al., 2009)
- Cross-correlation between observed and synthetic waveforms for time shift evaluation (obtained through wave equation solvers: expensive!)

Time delay will be frequency-dependent: for example, 0.6 s time shift between P time in band [2-0.5 Hz] and P time in band (0.1-0.03 Hz) …

- Starting model should be phase-compatible (could be cycle-skipped)

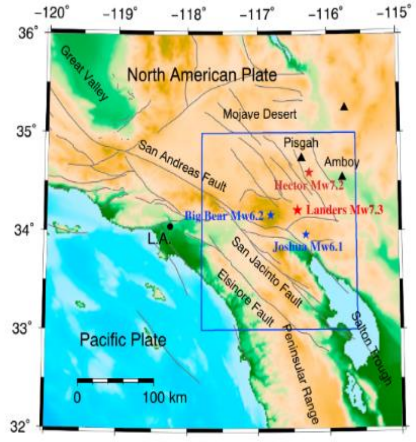
Interference between phases in the same window – non-linear effects (Nolet, 2008)

Model « see » the interference and not each phase … in a dynamic way (updated phase delays)

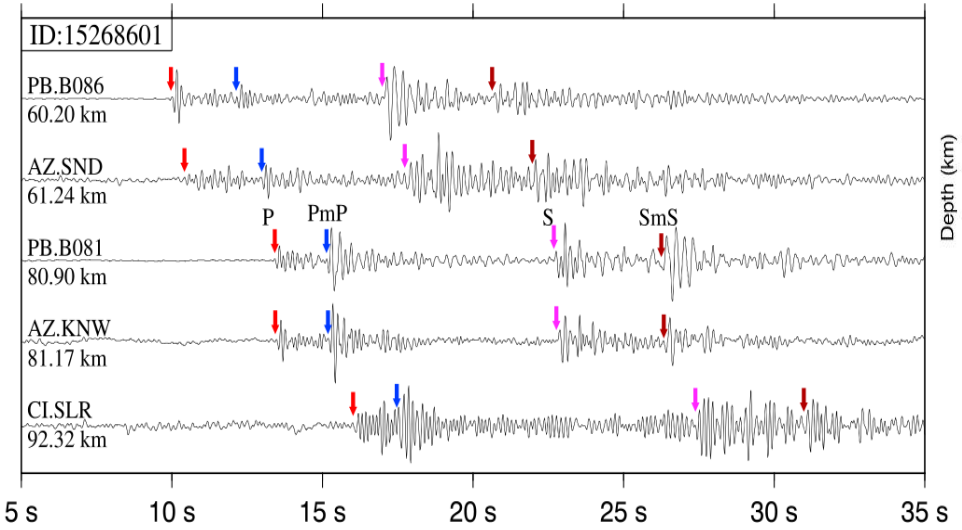
Wave-equation tomography

Wave Equation Tomography (WET)

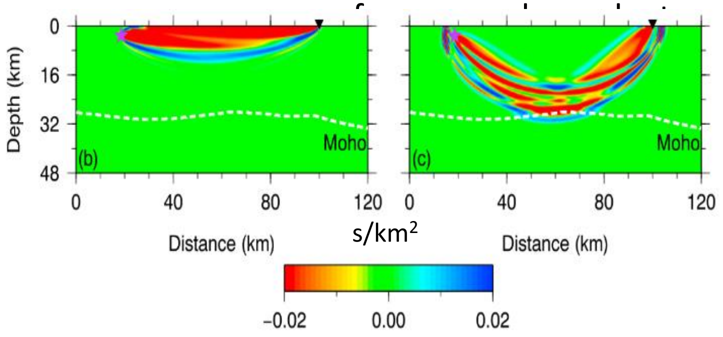
(more and more used tool in seismology)



(Huang et al, 2016)



Wavepaths: first-Fresnel zone



Time-domain cross-correlation inside a given window: « traveltime » sensitive to frequency

- Sensitivity kernel from Wave Equation (WE) (Tape et al., 2009, 2010; …)

- Sensitivity kernel from Ray Theory (RT) (Dahlen et al, 2000; Dahlen & Nolet, 2005)

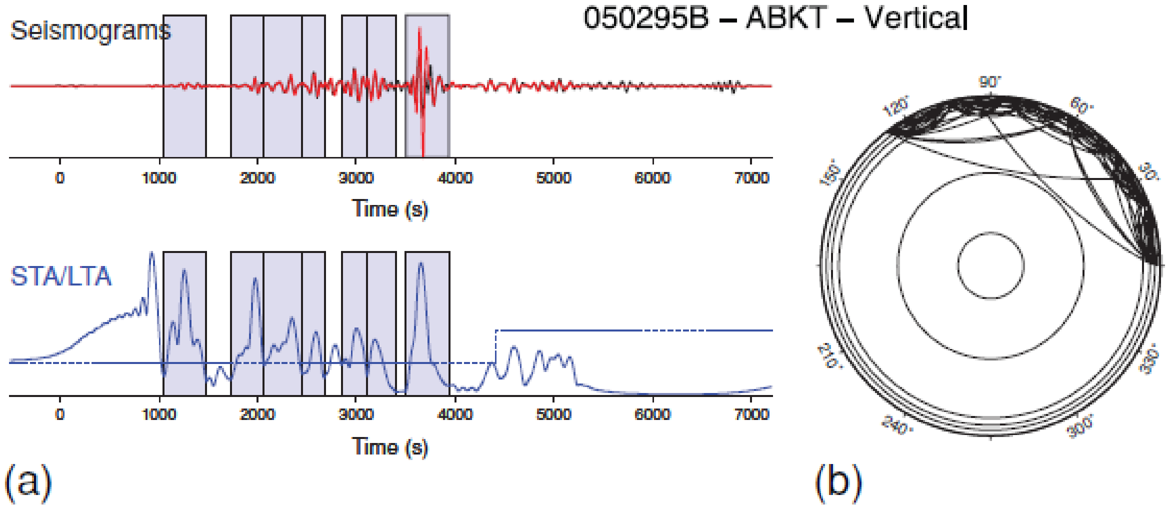
Sensitive to both amplitudes

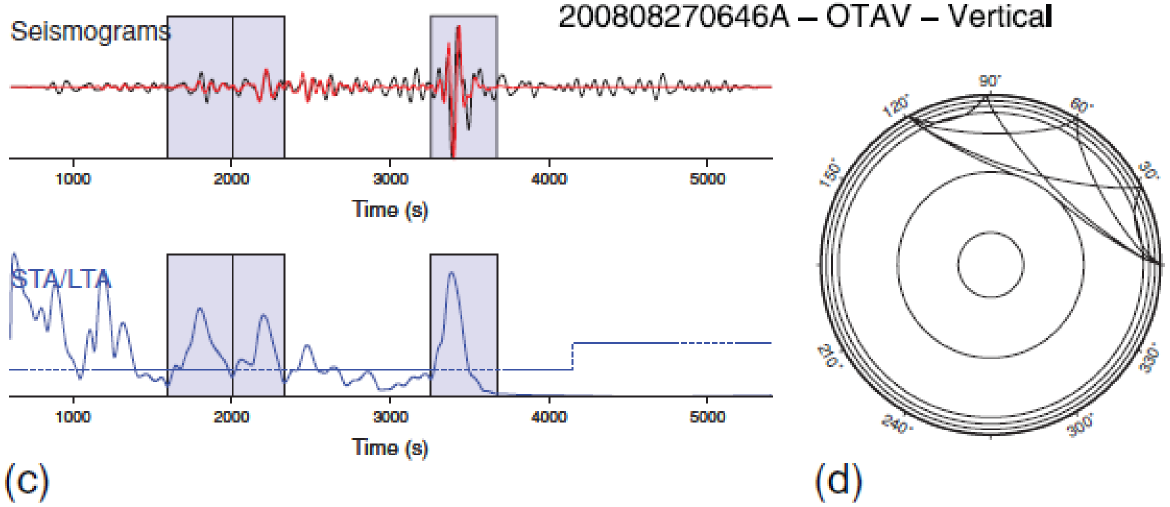
Sensitive to synthetic amplitudes

Consistency between finite-frequency data and ray concept?

Born approximation (Lippmann & Schwinger, 1950; Dahlen & Tromp, 1998; Zhao et al., 2000; Zhao & Chevrot, 2011a,b)

Window definition: Xcorrelation





Cross-correlation
between observed and synthetic seismograms

Starting with an already quite precise prediction of seismograms for amplitude comparison

(LTA/STA;
Maggi et al, 2009)

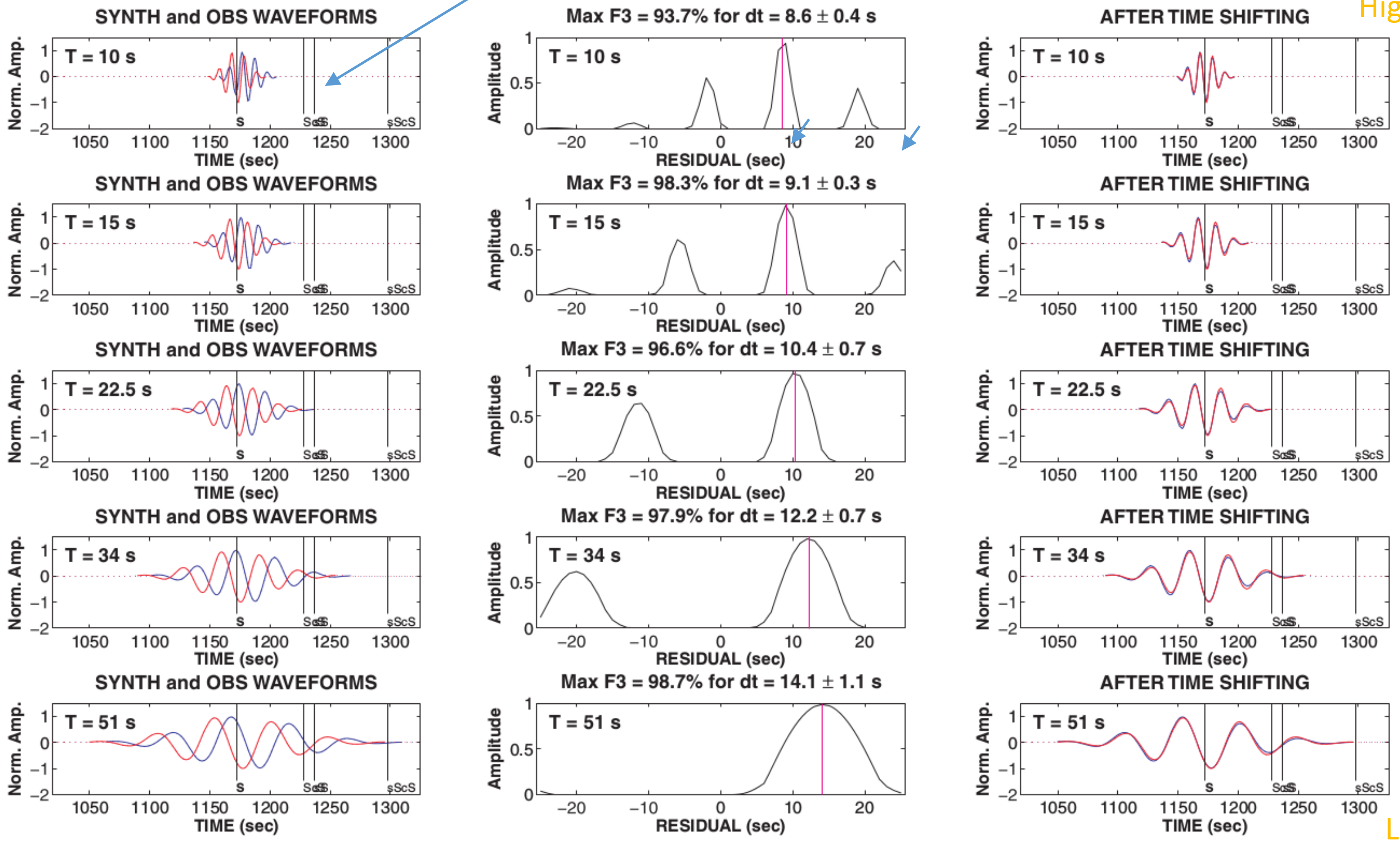
(Wavelet Freq/time;
Lee and Chen, 2013)

Dynamic phase measurement: Xcorrelation

« Tagging »

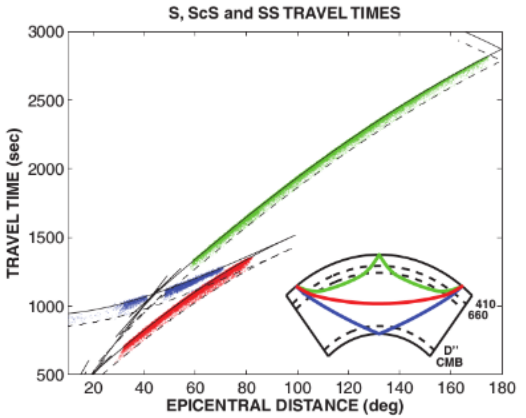
Interference btw ScS and sS

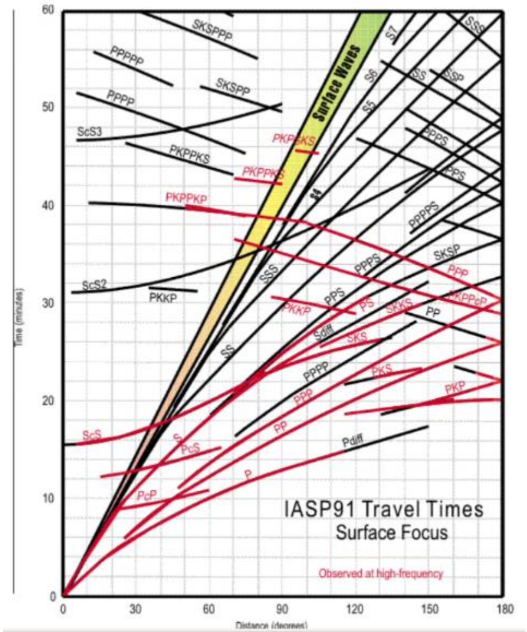
High frequency



Low frequency

Instantaneous phase

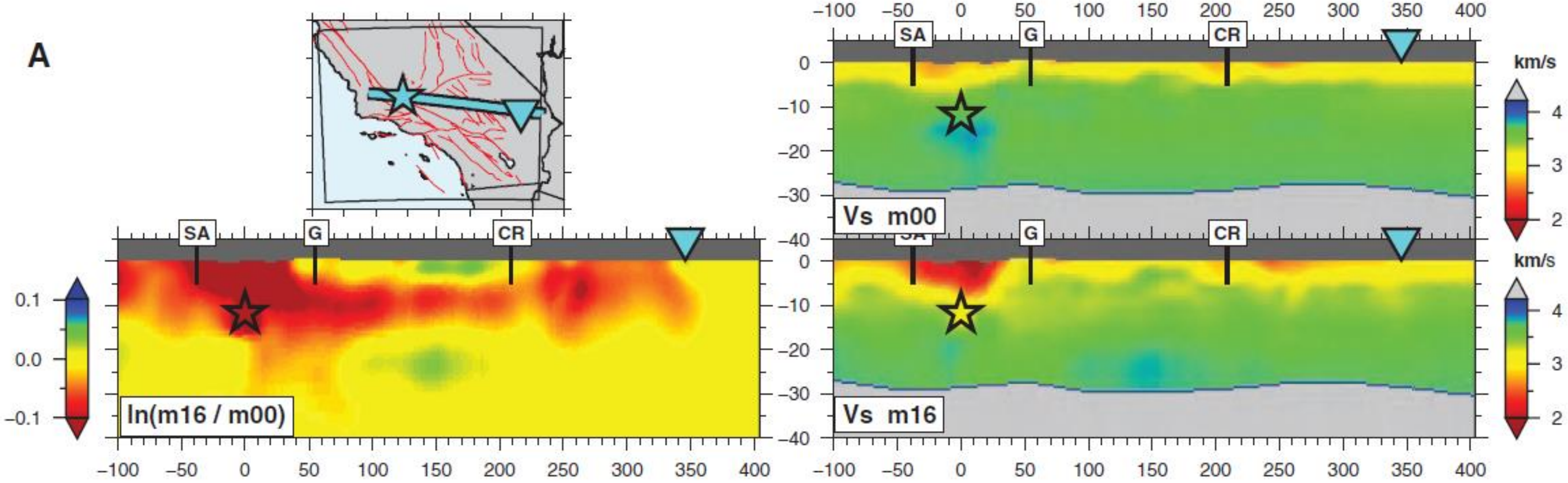




Zaroli et al (2010)

~400 000 phases S, SS, ScS « nearly » automatically

Wave-equation tomography



SCEC integrated initial model

New model at iteration 16

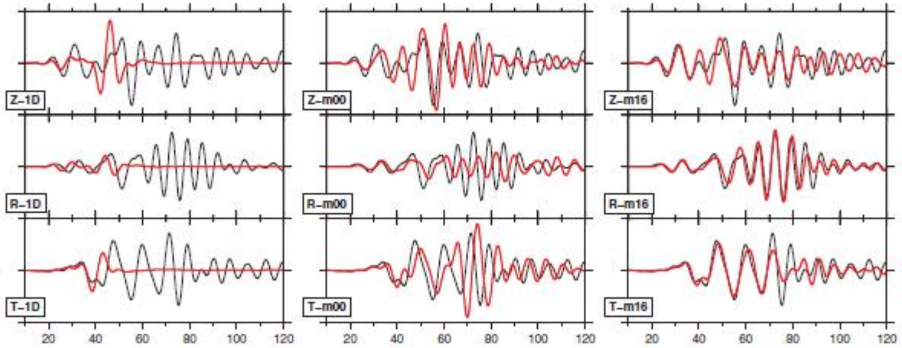
Earthquake database

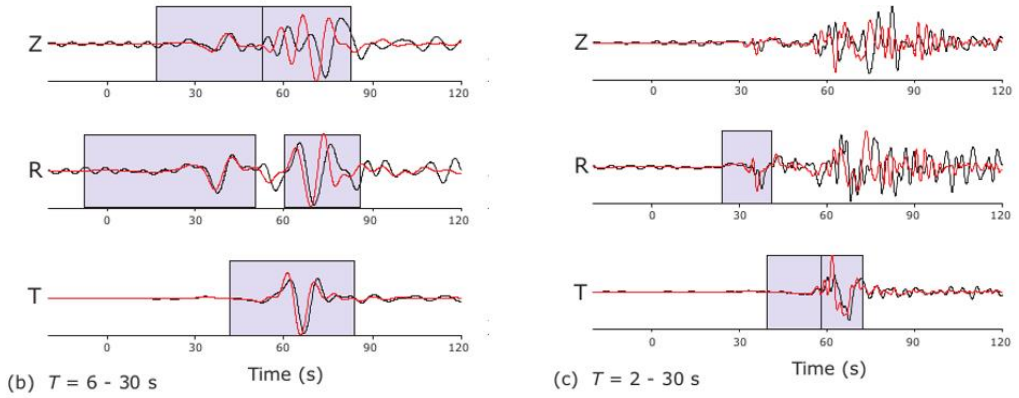
Phase delay through cross-correlation picks between observed and synthetic waveforms

vertical

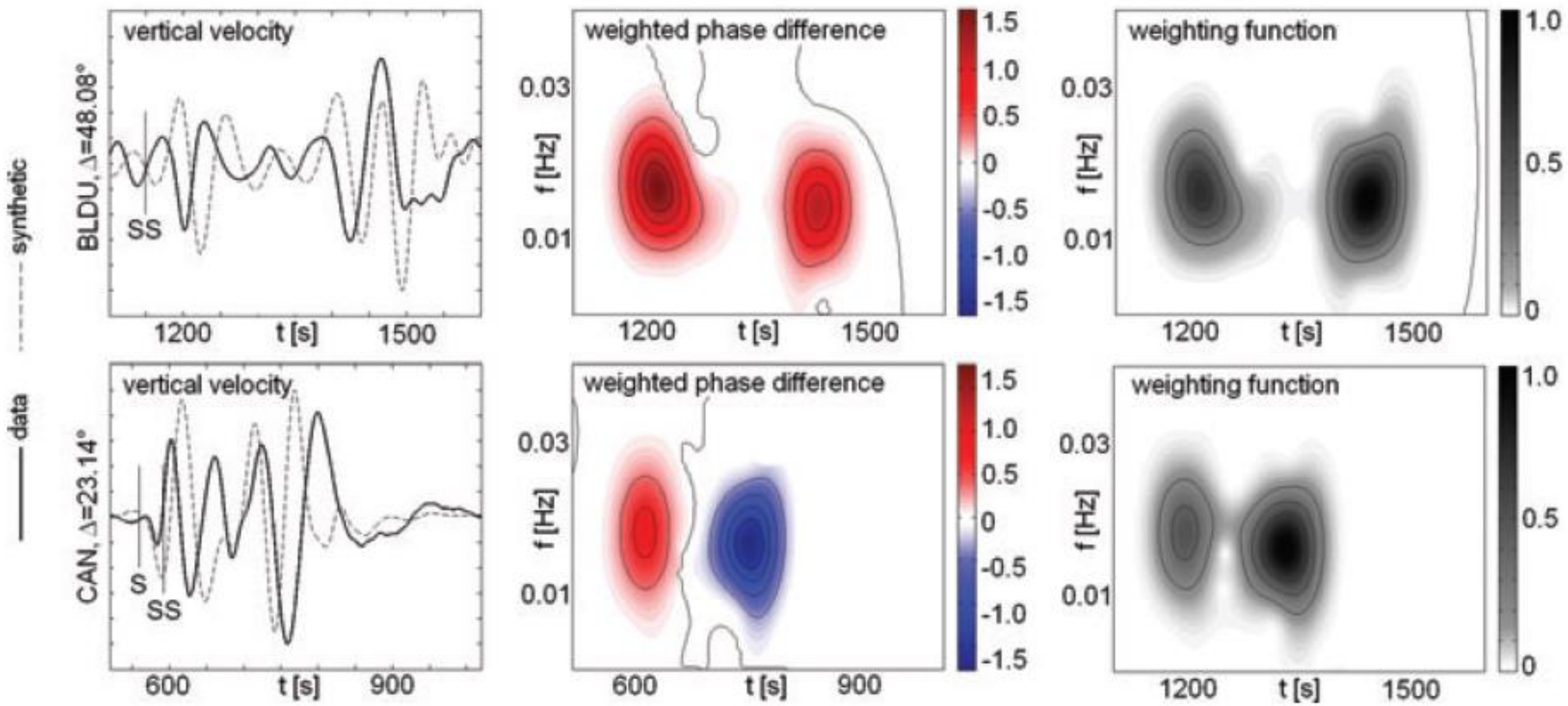
radial

transverse





WET: instantaneous phase misfit



(Fichtner et al., 2008; Fichtner et al., 2009)

Phases are not picked as for DRT or DET (unless dynamic wrapping)!

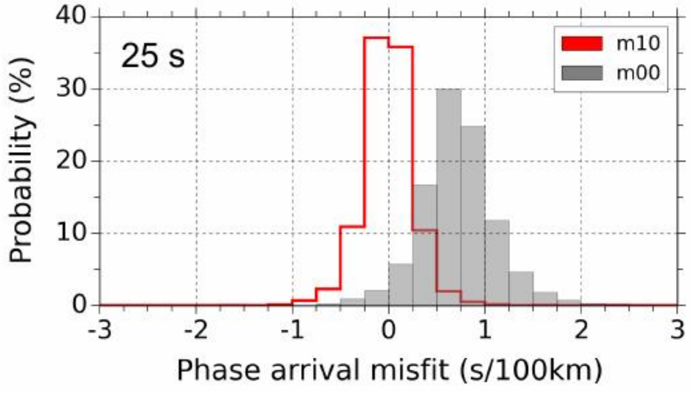
Phase delays are the extracted observables highly correlated to current model (Nolet, 2008)

Wave-equation tomography (ambient noise)

Real example for Europe using phases analysis and elastodynamic wave propagation

Lu et al (2018)

From ambient noise analysis

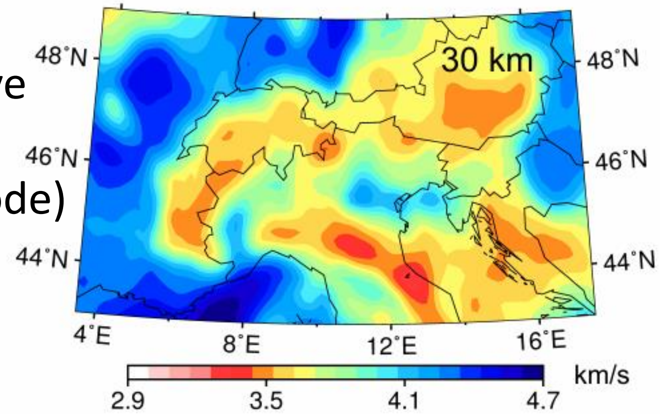


Seismology ☺

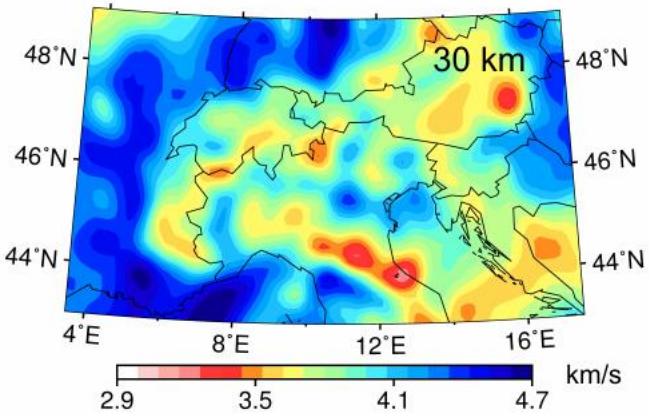
Which phases in WET
When doing ambient noise?

Dispersion curve analysis (fundamental mode)

Initial phase tomography



Wave equation tomography



Diffraction effects are included through wave equation propagation with finite frequency
Interference makes life difficult: all phases should be included when Xcorr (Nolet, 2008)

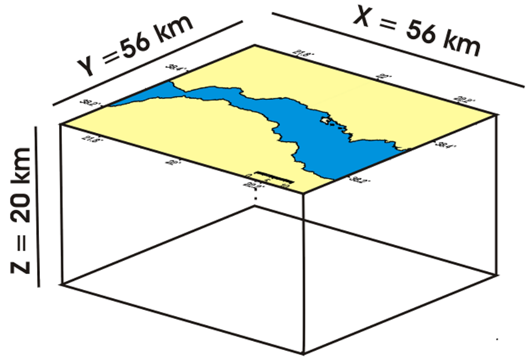
From active acquisitions

Seismic ☹/☺

Outline on first-arrival traveltime tomography

- Images at very different scales
- Waves and Phases: various concepts
- Few points on first-break ray-based tomography
- Illustration on 30-years Western Alps tomography
- First-break eikonal-based tomography
- First-break wave-equation-based tomography
- Hypocenter-velocity joint inversion
- Conclusion

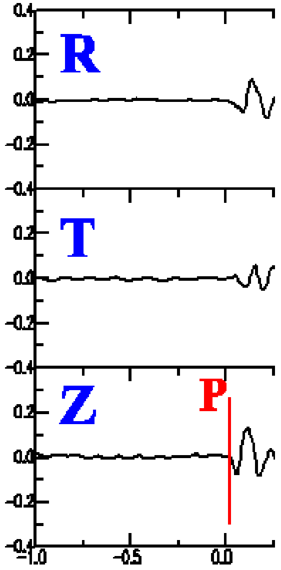
Hypocenter-velocity inversion



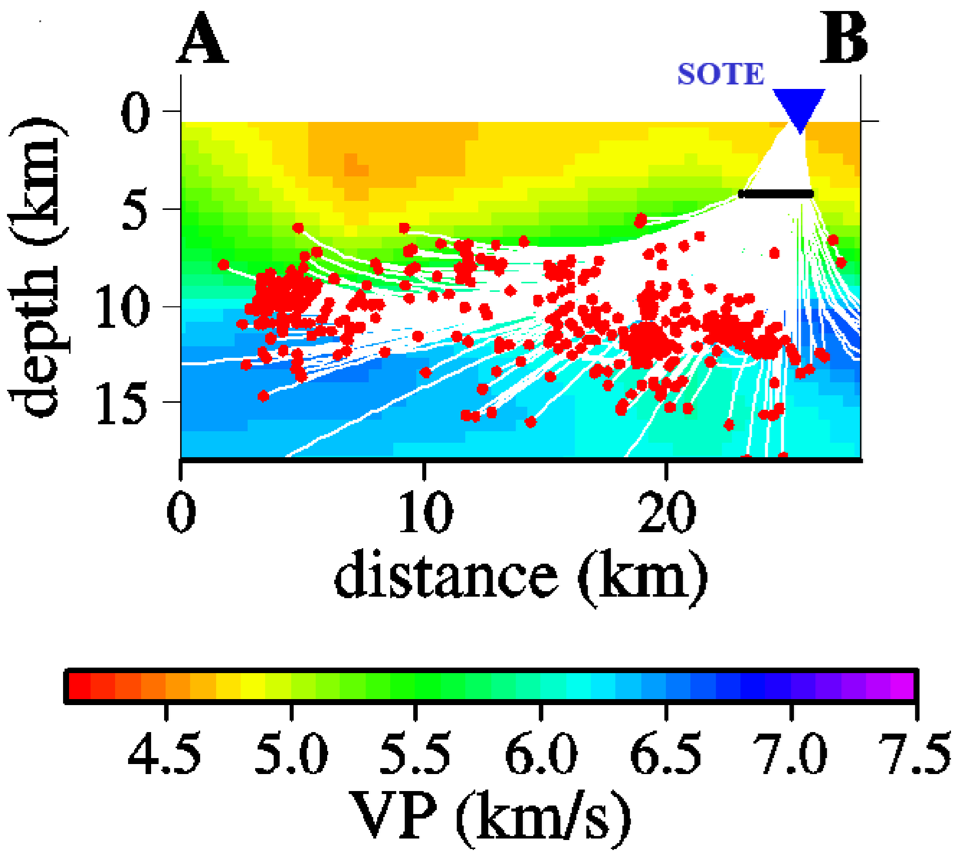
Target zone

Latorre et al. (2004)

Data: picked times



on seismograms



- Reconstruction at the local scale

Hypocenter parameters: $\mathrm{h} = (x_1^h, x_2^h, x_3^h, \tau_0^h)$

Velocity structure: $v(x_1, x_2, x_3)$ (slowness u)

- Input time data (~ 40 stations in 90's)

10000-50000 (P and/or S) picked times

- Discretization for local targets

1000-5000 events: few 1000 unknowns

Velocity/slowness values on grid 10x10x10

5.6 km x 5.6 km x 2 km: few 1000 unknowns

Modern seismic network
500-1000 stations are standard
worth revisiting such an inversion

Multi-parametric inversion

Travel time inversion is a non-linear problem which is solved through linearized system

$$r_k^i = (t^{obs} - t^{syn}) = \sum_{l=1}^{3} \frac{\partial T_k^i}{\partial x_l^i} \delta x_l^i + \delta \tau_0^i + \int_i^k \delta u \, dl$$

where time residual r_k^i from event i at station k

and integral $\int_i^k \delta s dl$ is computed along the ray from event to station

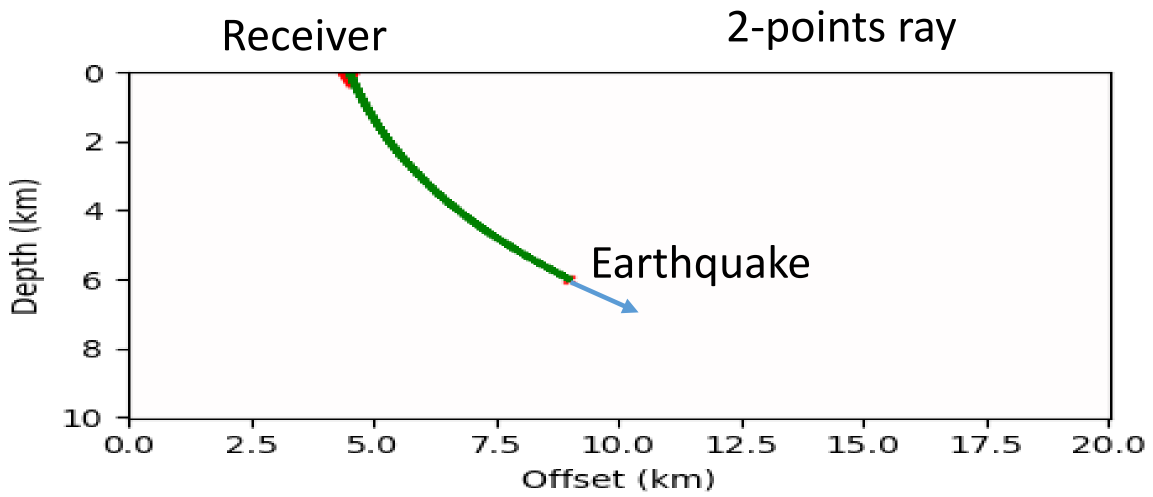
Discrete formulation of the slowness u

$$r_k^i = \sum_{l=1}^{3} \frac{\partial T_k^i}{\partial x_l^i} \delta x_l^i + \delta \tau_0^i + \sum_{j=1}^{m} \delta u_j \frac{\partial T_k^i}{\partial u_l^{i,k}}$$

Wanted parameters $\delta x_l^i, \delta \tau_0^i, \delta u_j$

Linear system at each hypocenter/slowness update

$$\frac{\partial T}{\partial h} dh + \frac{\partial T}{\partial u} du = r$$



Multi-parametric inversion

Travel time inversion is a non-linear problem solved through linearized system

Linear system at each hypocenter/slowness update

$$\frac{\partial T}{\partial h} dh + \frac{\partial T}{\partial u} du = r$$

All unknowns have to be inverted all at once

☹ Multi-parameter inversion
(cross-talk difficult to mitigate)

One can try an alternate inversion by locating first quakes in the current model and then by inverting slowness while keeping events fixed.

As shown by Thurber (1992) and Roecker et al. (2006), slowness perturbation will minimize any changes in hypocenter locations.

Annealing (SVD) strategies proposed by Pavlis & Booker (1980), Spencer & Gubbins (1980) or Rodi et al (1981) do not behave correctly.

Scaling & normalization are required: many numerical trials are needed

Three possible strategies

- Alternating strategy (Stork a Clayton, 1986)

$$\underset{h}{\operatorname{argmin}}\left\|d(u,h)-d^{obs}\right\|_2^2 \quad \fallingdotseq \quad \underset{u}{\operatorname{argmin}}\left\|d(u,h)-d^{obs}\right\|_2^2$$

Strong cross-talk

- Joint inversion (Bishop et al., 1985)

$$\underset{u,h}{\operatorname{argmin}}\left\|d(u,h)-d^{obs}\right\|_2^2$$

Scaling model parameter and Fréchet derivative

- Variable projection (Golub & Pereyra, 2003)

Consistency (Chauris et al., 2002; Guillaume et al., 2008)

$$\underset{u}{\operatorname{argmin}}\left\|d(u,h(u))-d^{obs}\right\|_2^2$$

Subproblem to be performed

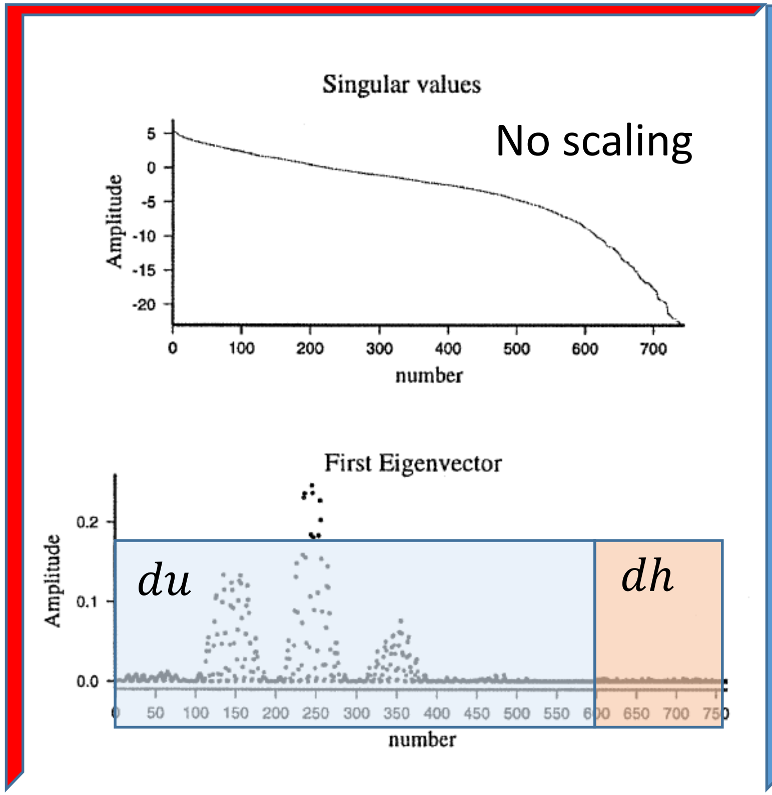
Different sensitivities

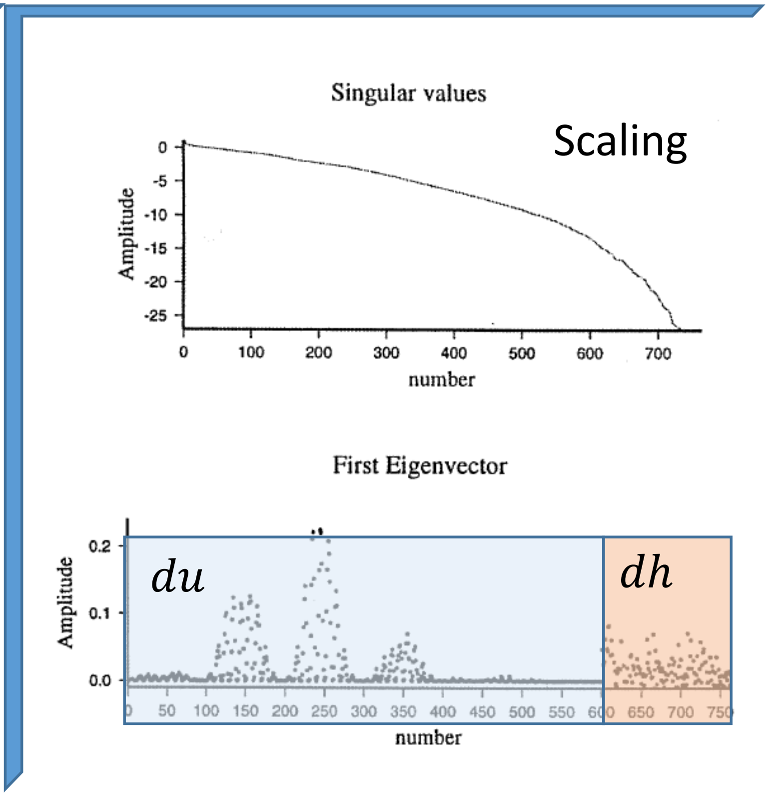
$$\frac{\partial T}{\partial u}du + \frac{\partial T}{\partial h}dh = r = J\begin{bmatrix} du \\ dh \end{bmatrix}$$

SVD analysis of Jacobian J
(Lemeur et al, 1997)

All softwares include some kind of scaling/normalization between these # classes of parameters …

Aside potential slow convergence, experience is required from users





Could we do it better?
Yes, we can …

$$\left(\frac{\partial T}{\partial u} + \frac{\partial T}{\partial h}\frac{\partial h}{\partial u}\right)du = r = \frac{dT}{du}du$$

Single class: slowness values

One hypocenter hypothesis?

Does it make sense to locate a pointsize quake in an inaccurate velocity structure?

Do split the quake database into few smaller databases enabling the location of virtual hypocenters: instead of one hypocenter for an earthquake, we have as many as the number of database subsets.

Updating the velocity structure such as these virtual hypocenters move to an unique position.

Question: is it a feasable strategy?

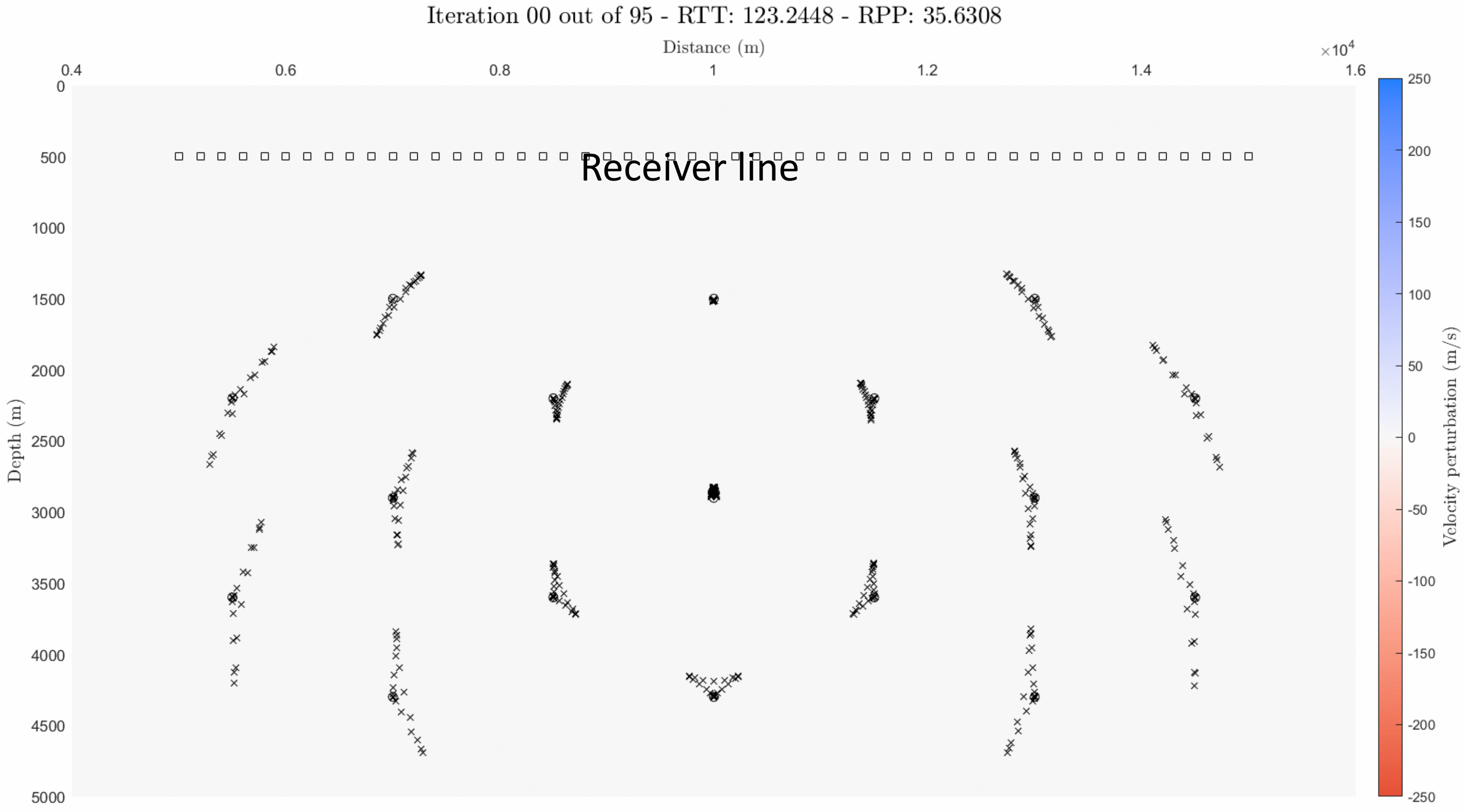
Maybe when considering similar concept from slope tomography **(Sambolian et al., 2021)**

Earthquake associated with many picked times

Simple example: event focusing

Circular velocity anomaly (250 m/s)

Earthquakes are recorded at all stations



In principle, it may work!!!

Outline on first-arrival traveltime tomography

- Images at very different scales
- Waves and Phases: various concepts
- Few points on first-break ray-based tomography
- Illustration on 30-years Western Alps tomography
- First-break eikonal-based tomography
- First-break wave-equation-based tomography
- Hypocenter-velocity joint inversion
- Conclusion

Take-away message

<u>Method-oriented message</u>

- Delayed ray-based and eikonal-based tomography **agnostic to frequency content**: user must drive these tools with this external information to be designed
  - Both approaches **same computational complexity**
  - Useful for interpretation **(trajectories)**
  - Hypocenter-velocity inversion
  - Useful for **uncertainty quantification after WET**

- Delayed wave-equation-based tomography is **sensitive to frequency content** and need a **fair prediction** of the synthetic waveform for doing cross-correlation between windowed traces
  - **Significant increase** in computational complexity
  - Hypocenter-velocity wave-equation based tomography **missing!**

Take-away message

<u>Application-oriented message</u>

Joint hypocenter-velocity inversion requires **adhoc sensitivity balance** between velocity parameters and hypocenter parameters: no consistent approach exists as far as I know.

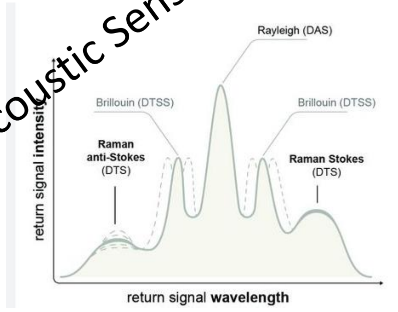
Initial model (and initial earthquake locations to lesser extent) design important!

Grid/Laplacian smoothing and **model gradient smoothing** play different roles

Increasing the **acquisition density** may sample better the near-surface zone

DAS is expected to be a game changer … in the near future???

Distributed Acoustic Sensing



End …

Questions?

Thank you very much
For your attention

Many thanks to sponsors of the SEISCOPE consortium

Fermat principle to Eikonal equation

First-arrival traveltimes follow Fermat principle of minimum time along any trajectory connecting the starting point and the end point.
The related variational problem can be written

$$\delta \int u(l)dl = 0$$

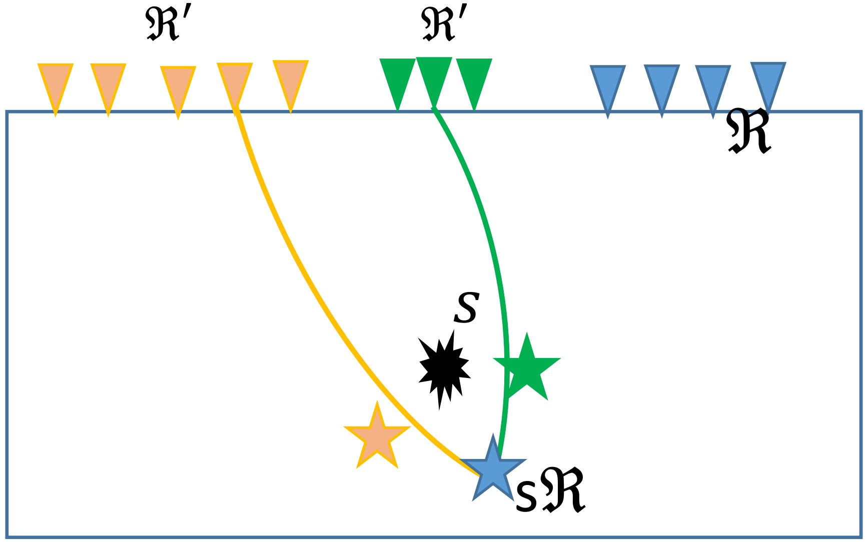
where $u(l)$ is the slowness at a given point and the curvilinear coordinate at this point is given by the quantity l.

The Eikonal equation can be obtained from the wave equation. It can also be thought of as the Hamilton-Jacobi equation of the above variational problem

(Kalaba, 1961 (isotropic case); Brandstatter, 1974 (anisotropic case)).

following slides are inspired from paper by Lakshminarayanan and Varadharajan (1997)

Virtual location parameters



For an event, many picks when dense acquisition

Subdivision of this dataset into subsets which induces a cosubset for each subset

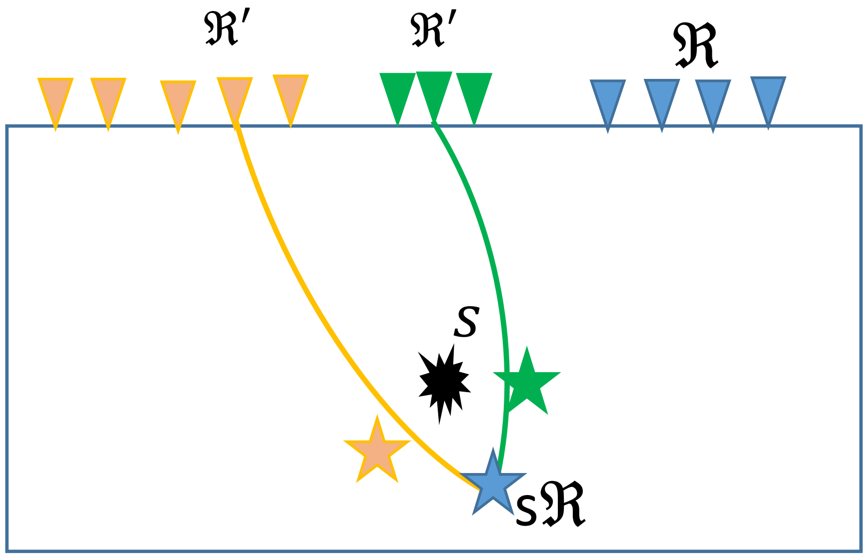
Locate virtual hypocenter for each subset

Evaluate residues at cosubset stations

True location S in black
Three virtual locations (blue, green, orange)
Residues of virtual blue location at blue and green stations (cosubset)

Adjoint variables of location parameters

Projection of location parameters
into adjoint variables
through misfit expression



Adjoint variables (α_{sr}) of the blue subset can be computed from residues at cosubset stations

$$\delta t_{sr,r'} = t_{sr,r'}(u) - t_{s,r'}^{obs}$$

$$\alpha_{sr}\partial_{x_{sr}} t_r(x_{sr}, z_{sr}) = \sum_{\mathfrak{R}', \mathfrak{R}' \neq \mathfrak{R}} \delta t_{sr,\mathfrak{R}'}\, \partial_{x_{sr}} t_{\mathfrak{R}'}(x_{sr}, z_{sr})$$

Eikonal PDE and adjoint PDE

Solving a non-linear PDE ☹

and a linear PDE ☺

(Isotropic case)

- ❑ Hamilton-Jacobi equation $\mathcal{H}(x, z, \nabla t_r(\mathrm{x}, \mathrm{z})) = \frac{1}{2}(\nabla t_r(x, z)^2 - u^2(x, z))$

  Forward mode

- ❑ Transport equation $\nabla.(\lambda_r(x, z)\nabla t_r(\mathrm{x}, \mathrm{z})) = \sum_s \sum_{r'=1, r' \neq r}^{N_r^S} (t_{sr',r} - t_{s,r}^{obs}) + \sum_s \alpha_{sr}$

  Reverse mode

  time residues

  Indirect contribution of location

  missing term when keeping earthquakes fixed

Two possible algorithms for each receiver

- ➢ Fast Marching Method (Vidale, 1988; Tsitsiklis, 1995)

Or

- ➢ Fast Sweeping Method (Tsai et al, 2003; Zhao, 2005) used by Sambolian et al (2019,2020)

$$r = \left(\frac{\partial T}{\partial u} + \frac{\partial T}{\partial h}\frac{\partial h}{\partial u}\right) du$$

Misfit gradient wrt slowness

Misfit gradient

$$\gamma_u(x,z) = \nabla_u \sum_s \phi_s(u)$$

$$\gamma_u(x,z) = \frac{1}{2}\sum_r u(x,z)\lambda_r(x,z)$$

Contribution from misfit

$$\sum_s \sum_{r'=1, r'\neq r}^{N_r^S} \left(t_{sr',r} - t_{s,r}^{obs}\right)$$

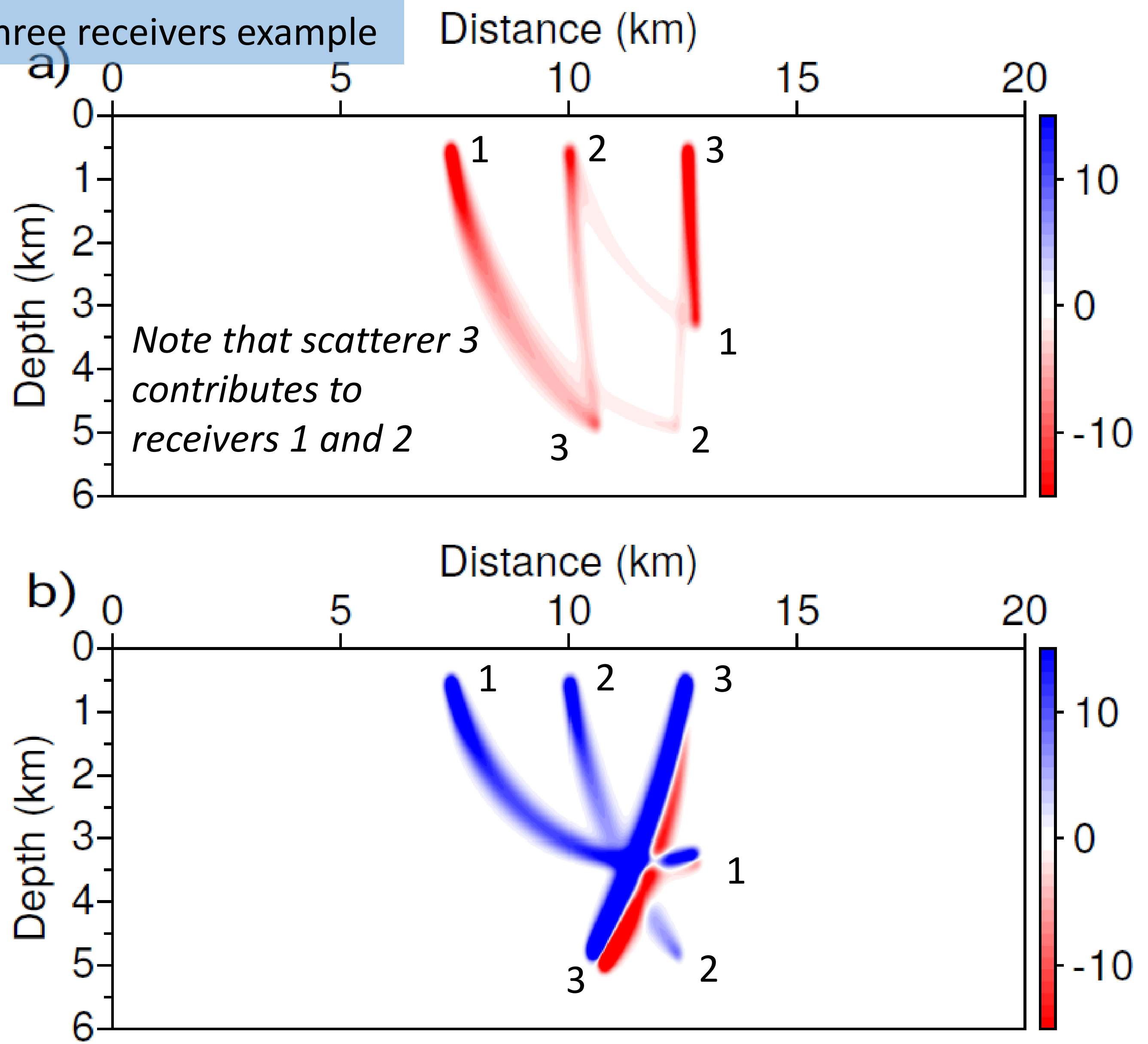
For each receiver r (1,2,3 ...)

Contribution from location

$$\sum_s \alpha_{sr}$$

In the alternate strategy of the standard formulation, the contribution of location is missing

Three receivers example



Eikonal solver – single solution

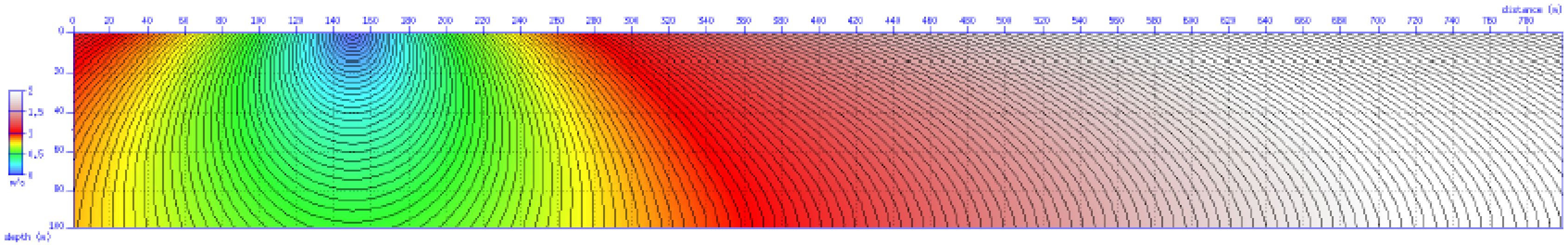
Two strategies with pros and cons

Fast marching method

Fast sweeping method

Eikonal solver – sweeping or marching ?

Synthetic example:



(c) Traveltime map

Forward problem: the eikonal equation

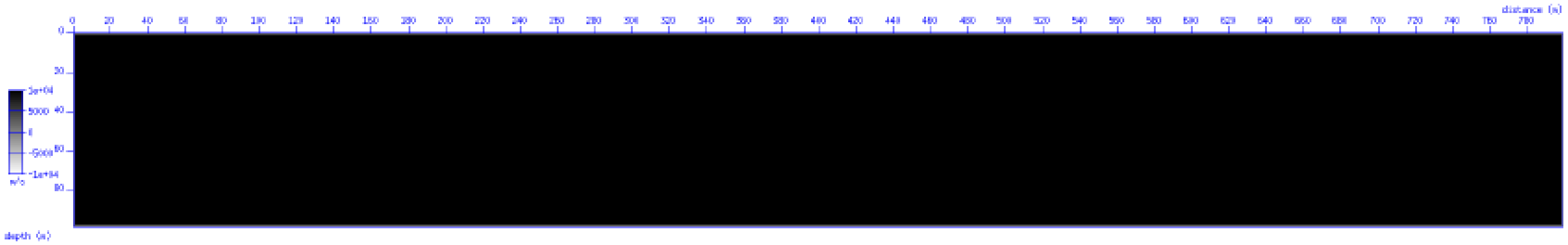
$$|\nabla\tau| \quad = \quad s$$

Good accuracy is needed to compute traveltime residuals (and raypaths).

[Podvin and Lecomte, 1991, Zhao, 2005, Li et al., 2008, Fomel et al., 2009, Belayouni, 2012]

(Noble et al, 2014)

Synthetic example:



(d) λ initial

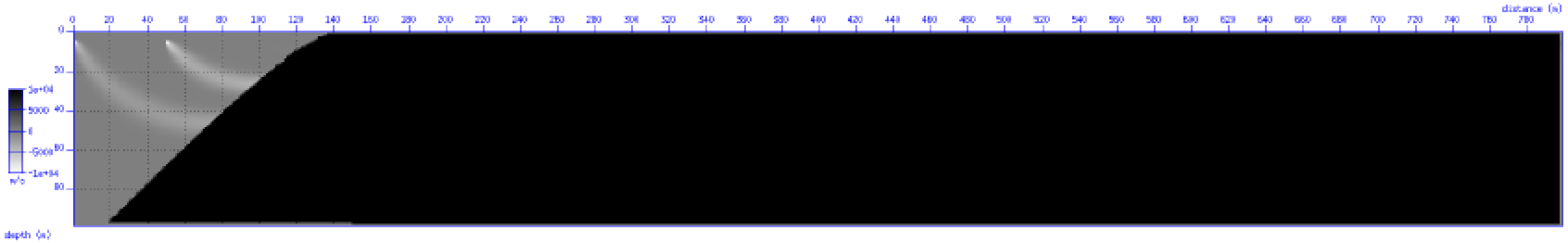
Adjoint problem: a transport equation

$$\nabla(\lambda \nabla \tau) \;=\; -\sum_r \mathcal{R}_r^t \, \delta t_r$$

We use Fast Sweeping Method.
[Zhao, 2005, Sei and Symes, 1994, Leung and Qian, 2006, Taillandier et al., 2009]

Adjoint solver – sweeping method

Synthetic example:



(e) λ sweep 1 : DOWN-RIGHT

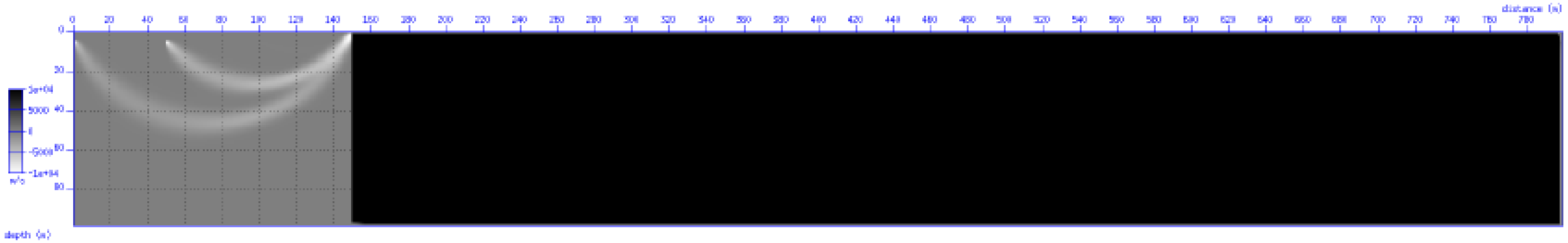
Adjoint problem: a transport equation

$$\nabla(\lambda\nabla\tau) = -\sum_r \mathcal{R}_r^t\,\delta t_r$$

We use Fast Sweeping Method.
[Zhao, 2005, Sei and Symes, 1994, Leung and Qian, 2006, Taillandier et al., 2009]

Adjoint solver – sweeping method solution

Synthetic example:



(f) λ sweep 2 : UP-RIGHT

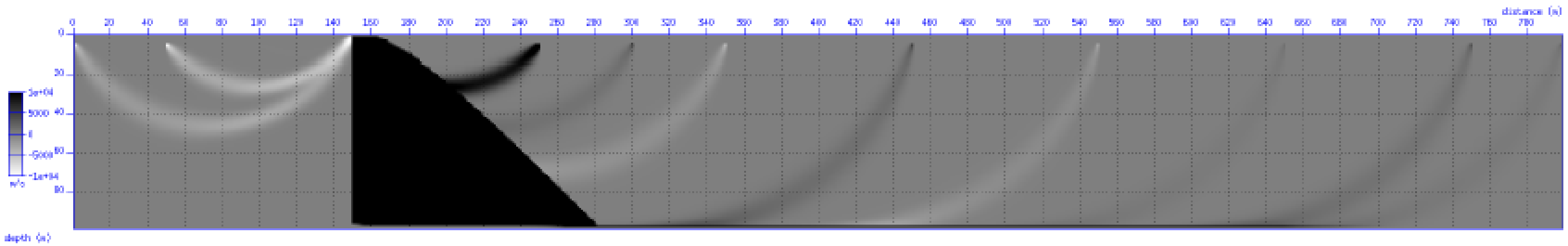
Adjoint problem: a transport equation

$$\nabla(\lambda \nabla \tau) \quad = \quad - \sum_r \mathcal{R}_r^t \, \delta t_r$$

We use Fast Sweeping Method.
[Zhao, 2005, Sei and Symes, 1994, Leung and Qian, 2006, Taillandier et al., 2009]

Adjoint solver – sweeping method

Synthetic example:



(g) λ sweep 3 : DOWN-LEFT

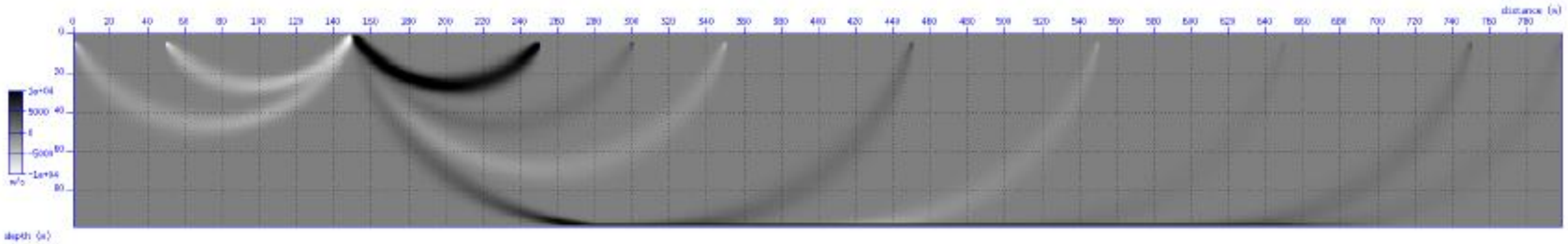
Adjoint problem: a transport equation

$$\nabla(\lambda \nabla \tau) = -\sum_r \mathcal{R}_r^t \, \delta t_r$$

We use Fast Sweeping Method.
[Zhao, 2005, Sei and Symes, 1994, Leung and Qian, 2006, Taillandier et al., 2009]

Adjoint solver – sweeping method

Synthetic example:



(h) λ sweep 4 : UP-LEFT

Adjoint problem: a transport equation

$$\nabla(\lambda\nabla\tau) = -\sum_r \mathcal{R}_r^t \, \delta t_r$$

We use Fast Sweeping Method.
[Zhao, 2005, Sei and Symes, 1994, Leung and Qian, 2006, Taillandier et al., 2009]

Eikonal Solvers

Fast marching method (FMM)

In a 2D layered medium

(tracking interface/wavefront motion : curve and surface evolution)

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

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

Let us assume T is known at a level z=cte

a) $\frac{\partial T}{\partial x}$

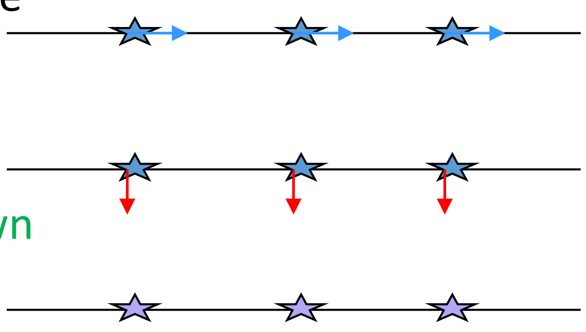
z=cte

b) $\frac{\partial T}{\partial z}$

z

c) down

z + dz



Compute $\partial T/\partial x$ along z=cte by a finite difference approximation

Deduce $\partial T/\partial z$ from eikonal

Extend T estimation at depth z+dz

EIKONAL SOLVER

ONLY FIRST-ARRIVAL!

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

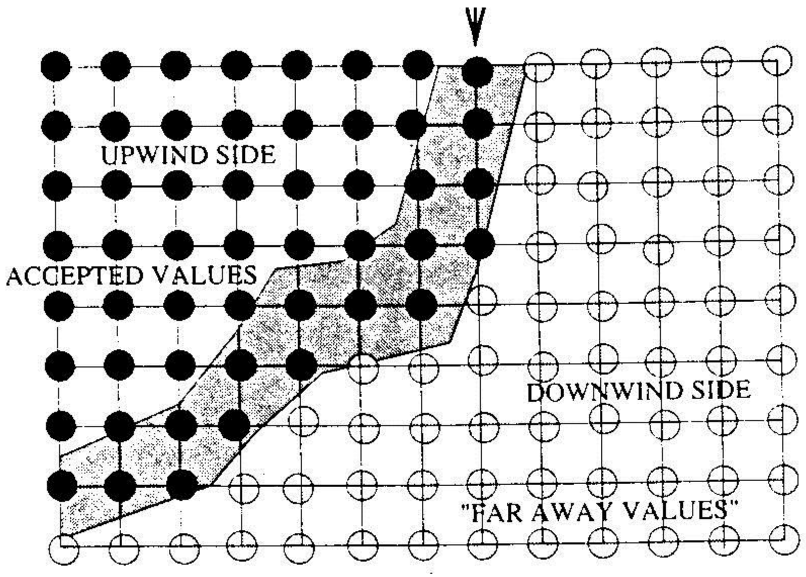
Fast marching method (FMM)

2D case

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

$$\frac{\partial T}{\partial z} = \pm\sqrt{\frac{1}{c(x,z)^2} - \left(\frac{\partial T}{\partial x}\right)^2}$$

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



From Sethian & Podovici (1999)

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

$$\mathcal{O}(N\ LogN)$$

where N is the number of grid points in a direction

The ENO or WENO stencil

How to estimate the discrete $|\nabla T|$?

A first-order Godunov upwind difference scheme

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

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

See Sethian’s book (1999)

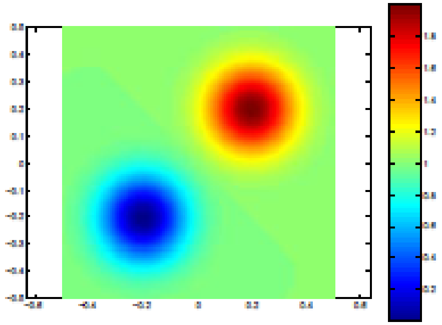
Fast sweeping method (FSM)

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

Computing T(x,z) is a stationay boundary problem

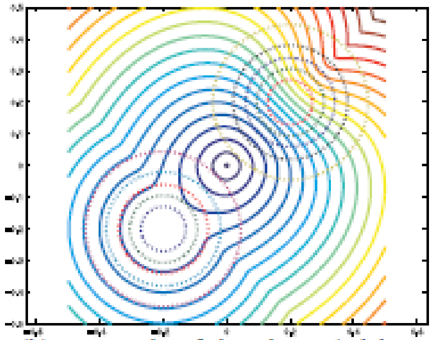
Gauss-Seidel iterations with alternating direction sweepings are incorporated into same upwind finite difference stencil: complexity in $\mathcal{O}(N)$.

Velocity



(a) graph of $f(x, y)$

Travel time



(b) contour plot of the solution (solid line) and $f(x, y)$ (dashed line)

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)

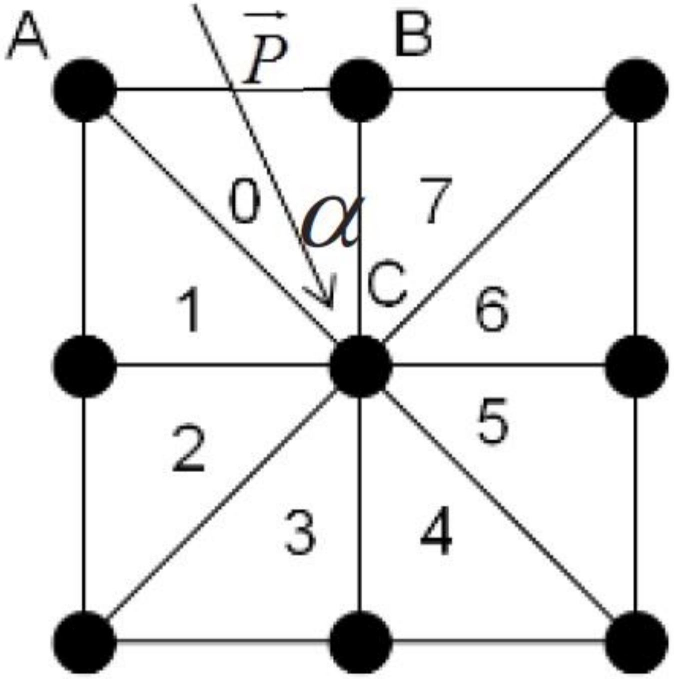
2D case

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

Any mesh could be used

$$T_C = T_A + \frac{\vec{AC}.\vec{p}}{AC}\mathrm{AC} \qquad T_C > T_A$$

$$T_C = T_B + \frac{\vec{BC}.\vec{p}}{BC}\mathrm{BC} \qquad T_C > T_B$$



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

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

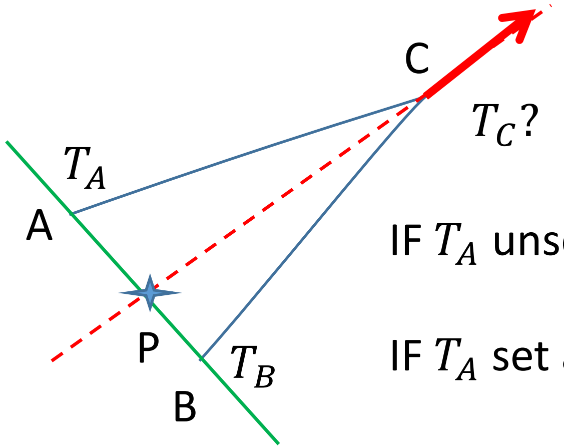
$$\begin{pmatrix} p_x \\ p_z \end{pmatrix} = \begin{pmatrix} MT_C + N \\ PT_C + Q \end{pmatrix} \Longrightarrow (MT_C + N)^2+(PT_C + Q)^2=\frac{1}{c^2(x,z)}$$

Han et al (2015)

Quadratic equation: real solution T_C needed!

Causality and viscosity

2D case



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

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

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

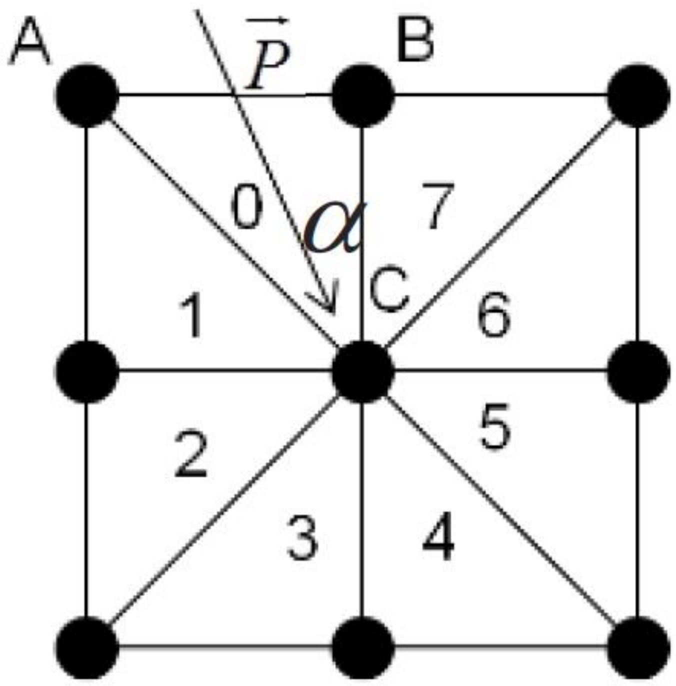
IF T_A unset and IF T_B set, compute T_C as if wave comes from B

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

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

Fast sweeping method (FSM)

2D case



We need to setup a set of values
which will be fixed:
at the source, $T = 0$, for example

Eight points stencil:

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

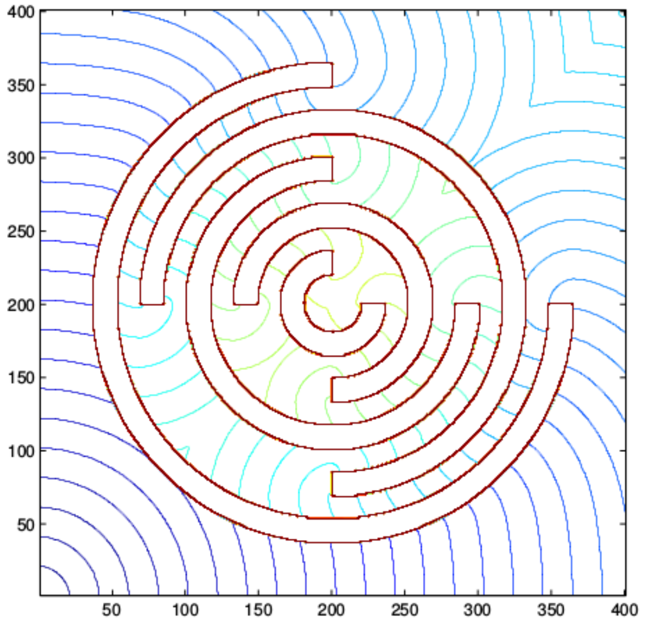
Sweeping technique:

Four sweeping when applying the stencil

Sweep 1: $i_i = 1, n_1, 1; i_2 = 1, n_2, 1$
Sweep 2: $i_i = n_1, 1, -1; i_2 = 1, n_2, 1$
Sweep 3: $i_i = n_1, 1, -1; i_2 = n_2, 1, -1$
Sweep 4: $i_i = 1, n_1, 1; i_2 = n_2, 1, -1$

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

Number of iterations



(c) distance to a point with five ring obstacle

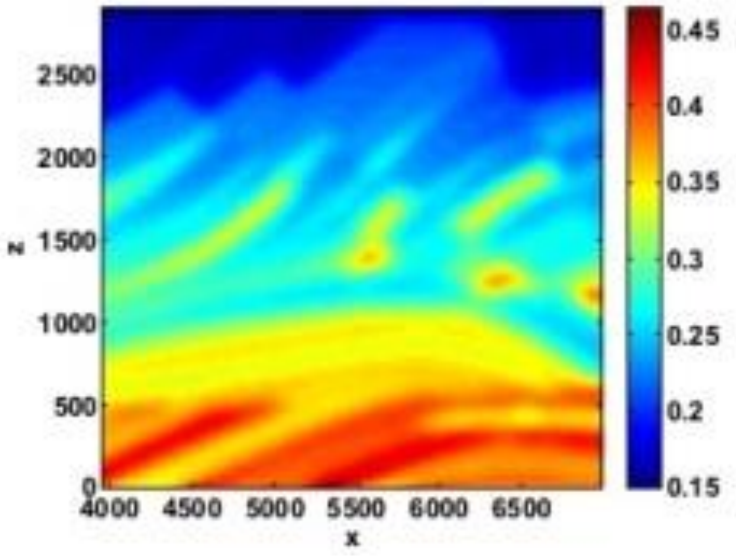
The number of iterations depend on the medium structure: one must be aware that the characteristics of the hyperbolic system should be sampled at least once by the sweeping loop.

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

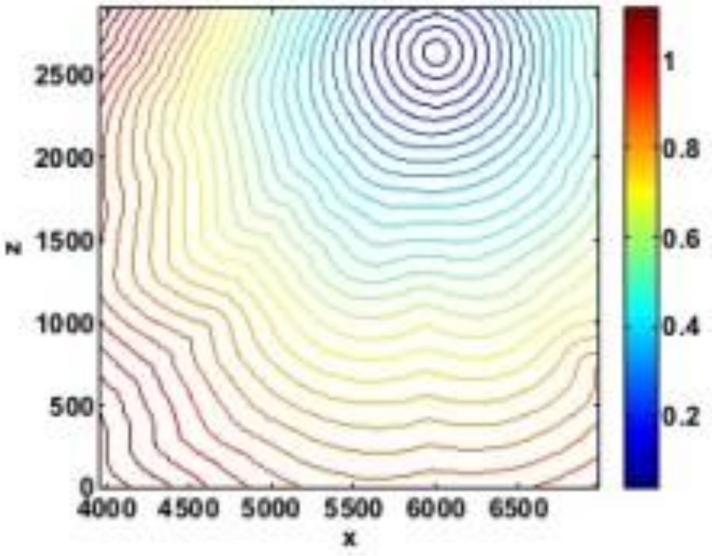
Few iterations are necessary to achieve convergence

2D examples

A smooth model constructed from Marmousi



Wavefronts as deduced from travel-time computation



(Luo and Qian, 2008)

Only first-arrival times

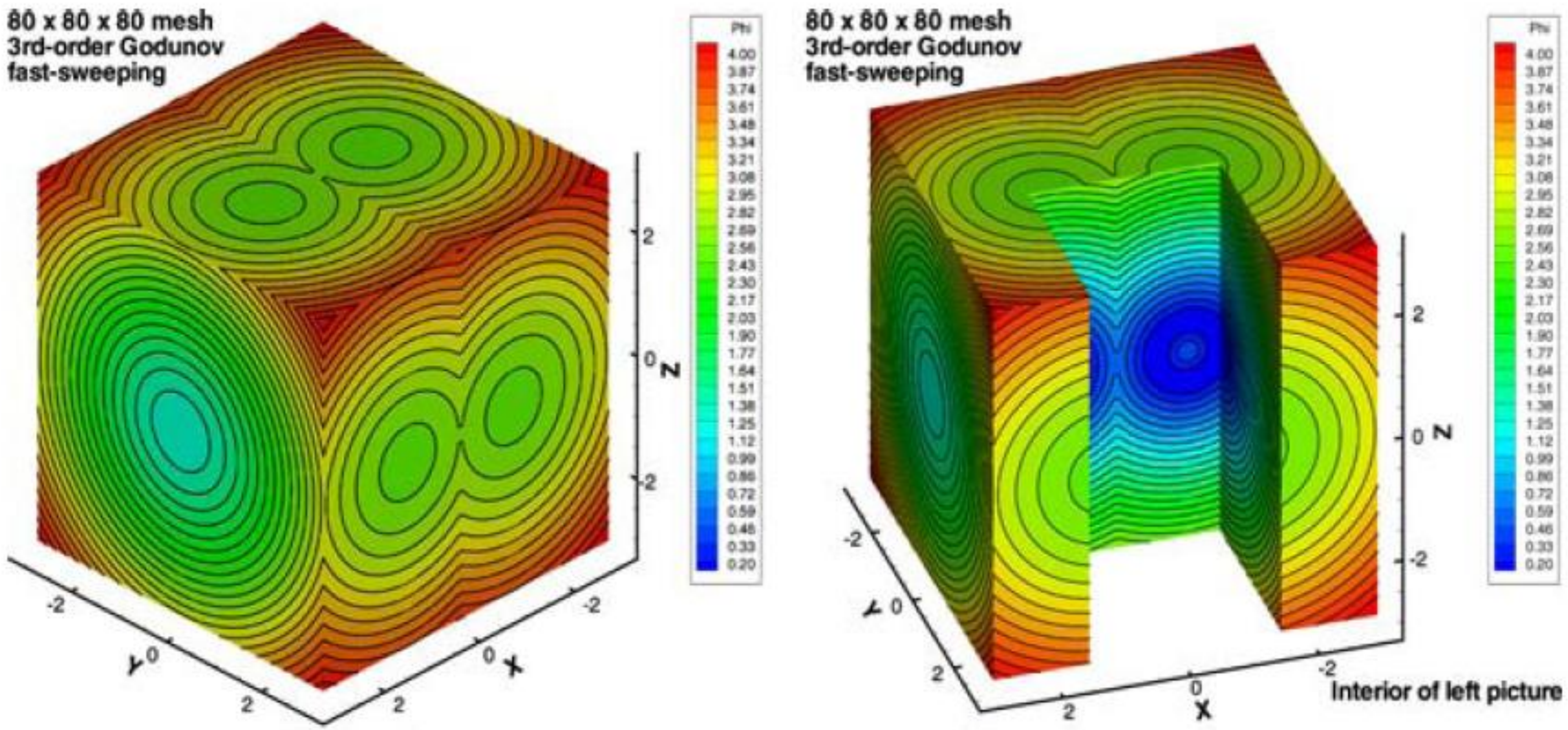
See Taillandier et al (1999) for an application to first-arrival time tomography using the adjoint formulation

3D example

$$\begin{cases} H(\phi_{x_1}, \ldots, \phi_{x_d}, x) = 0, & x \in \Omega \setminus \Gamma, \\ \phi(x) = g(x), & x \in \Gamma \subset \Omega, \end{cases}$$

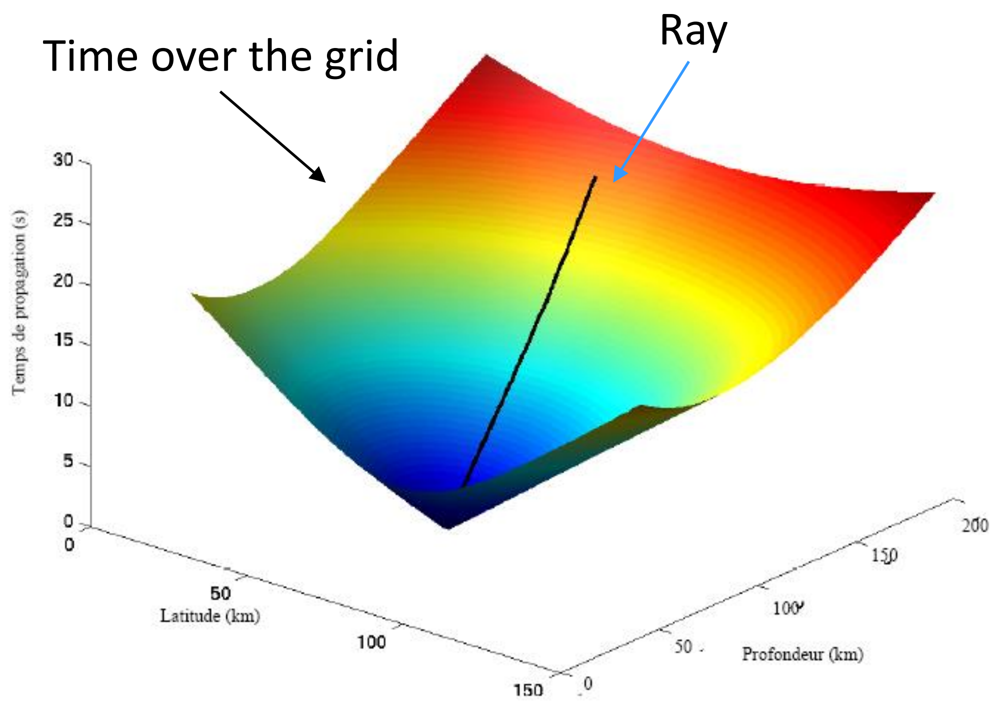
Zhang, Zhao and Qian (2005)

Static Hamilton–Jacobi Equations



Drawback: only first-arrival times!

Still very usefuf 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 {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?

Local ODE

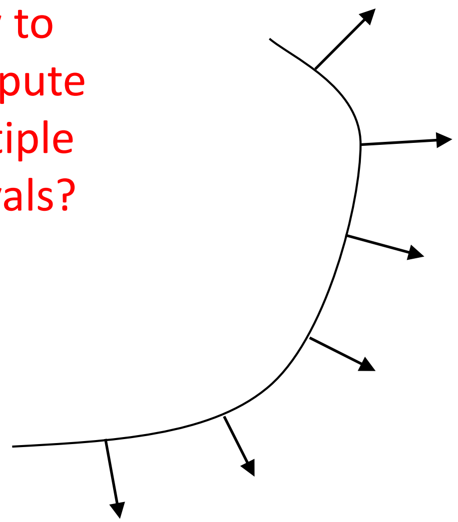
Semi-Lagrangian approach

Sampling control

Ray tracing by wavefronts

(Ray tracing by rays)

How to compute multiple arrivals?



Lambaré et al (1996)

Vinje et al (1993) widely used in NORSAR software

Evolution over time:

folding of the wavefront is allowed (still a significant curse of complexity!)

Dynamic sampling:

undersampling of ray fans

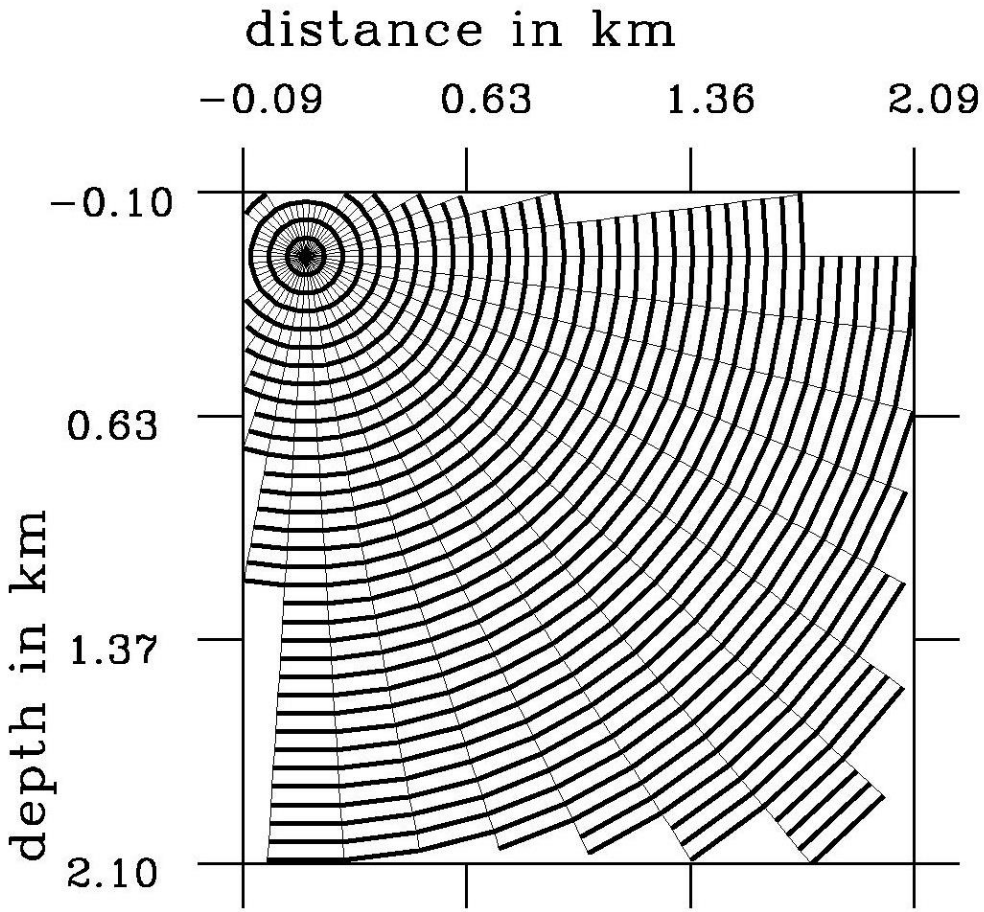
oversampling of ray fans

Keep an « uniform » sampling of the medium by rays

by tracking the surrounding density of rays

by estimating through paraxial approach the ray density

Ray tracing by wavefronts



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

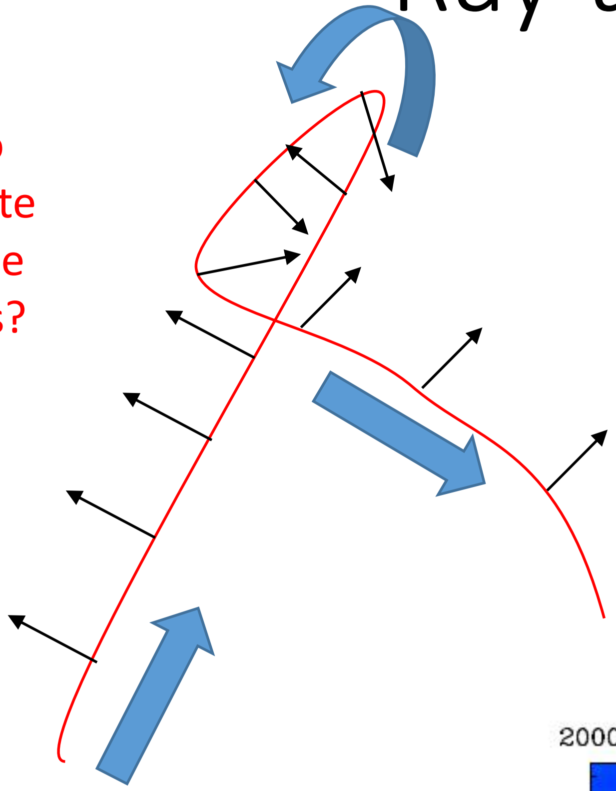
Keeping the sampling of the medium more or less uniform

Rays and wavefronts in an homogeneous medium.

(Lambaré et al., 1996)

Ray tracing by wavefronts

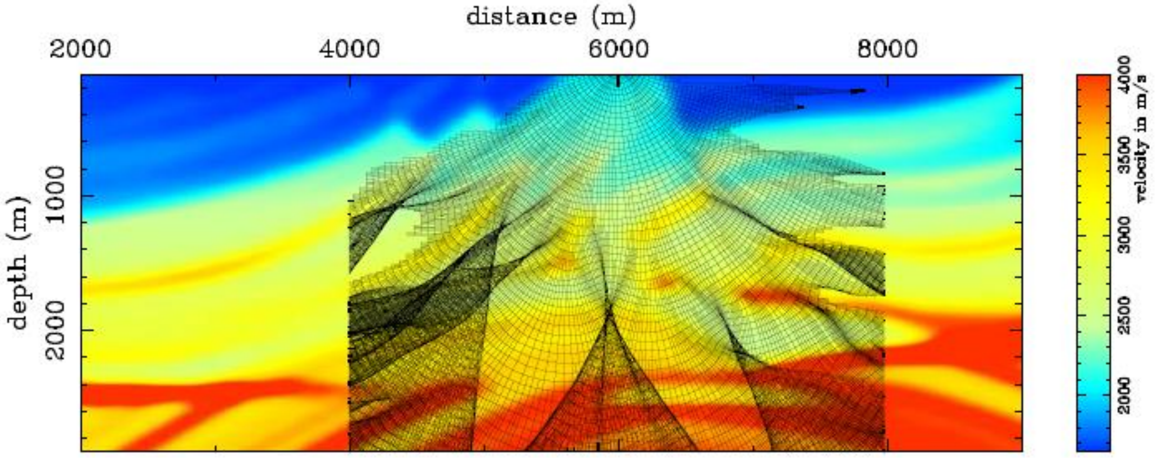
How to compute multiple arrivals?



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

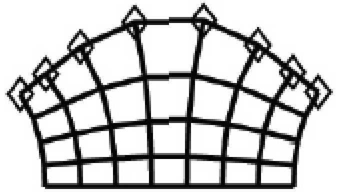


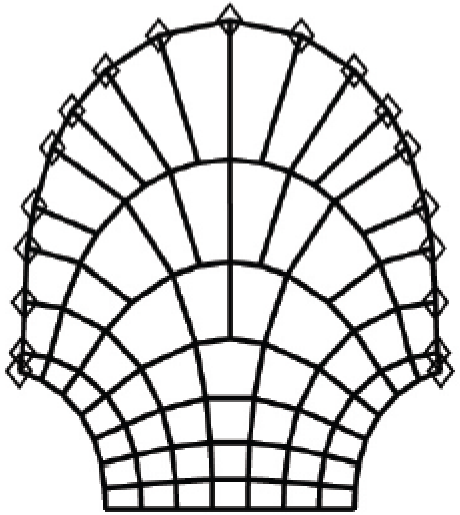
Smoothness is required !

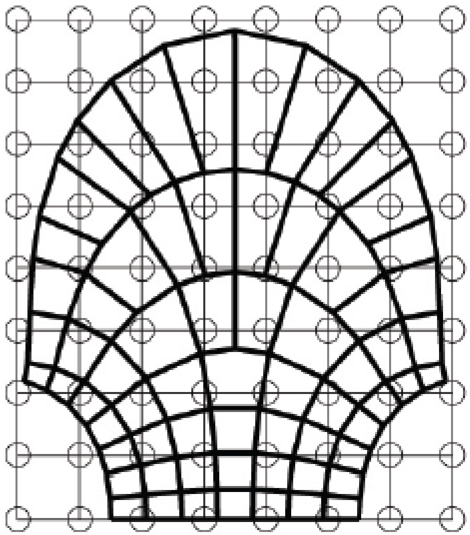
Wavefront construction [Vinje, Iversen, Gjöystdal, Lambaré, . . .]

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

(Runbord, 2007)







$\boldsymbol{x}(t_n, \alpha_j)$

Geometrical optics resolution

Do we need this complexity of tracking seismic wavefronts!

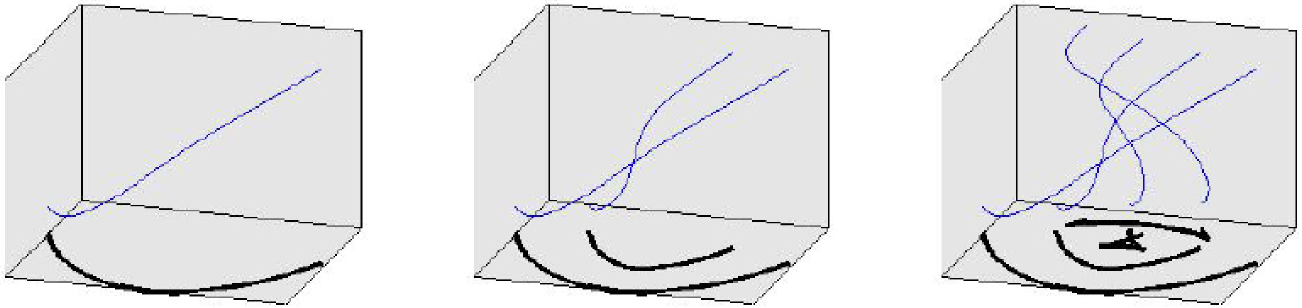
No scale as we are at infinite frequency !

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.

Phase space

Wavefront in *xy*-space typically non-smooth. Caustics appear.

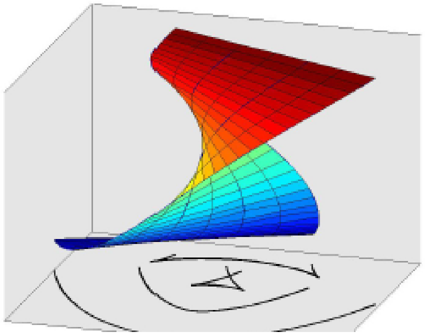


Introduce phase space $(\boldsymbol{x}, \boldsymbol{p})$, where $\boldsymbol{p} \in \mathbb{S}^{d-1}$ is local ray direction

Observation: Wavefront is a smooth curve in phase space.

- 2D problems: 1D curve in 3D phase space (x, y, θ).
- 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 $|\nabla_{\vec{x}} T(\vec{x})| = \frac{1}{c(\vec{x})}$ for sub-vertical rays and, therefore, we may choose the variable z as the evolution. We have

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

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 $\varphi\big(z, x(z), p_x(z)\big)$=0, i.e rays intersections at constant depth. Travel-times can be deduced by integration as well.

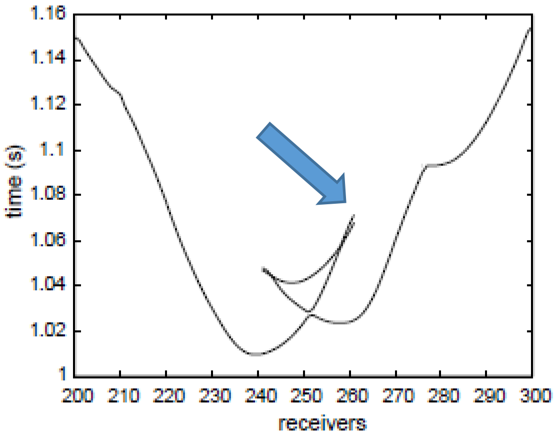
(Leung & Qian & Osher, 2004)

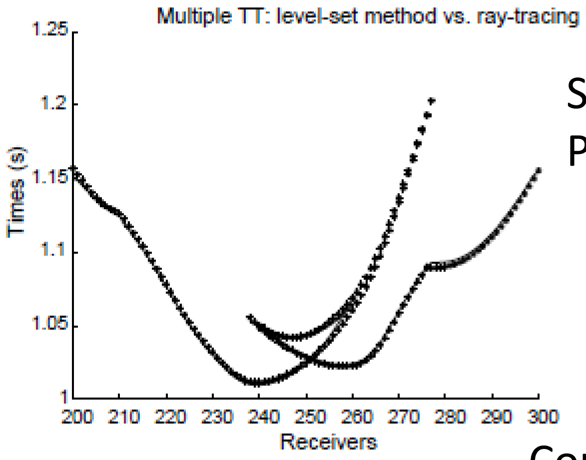
Application: Marmousi case

Medium from 4.8 km to 7.2 km in x and from 0 km to 3 km in x
Grid (x,z) = 384 x 122 with a stepping 25 m x 25 m
The source is at x=6.0 km and z=2.8 km

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

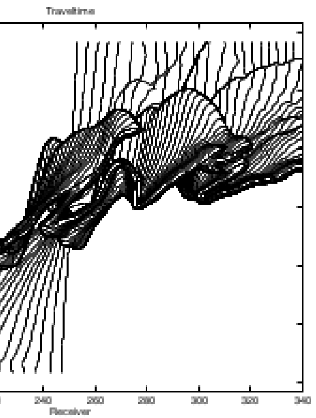
PDE

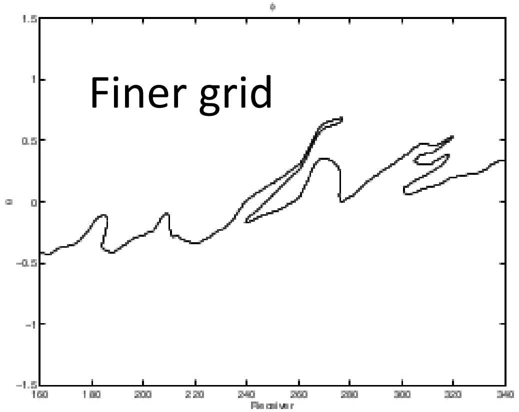


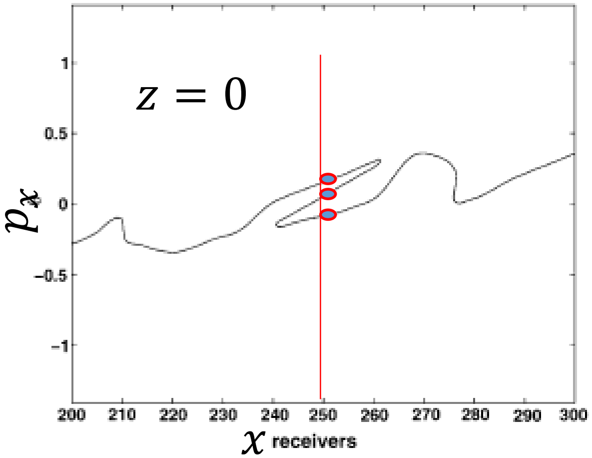


Superposition of ODE & PDE solutions

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







Contours for various z