Mathematics of Seismic Imaging
Part I

William W. Symes

PIMS, July 2005

A mathematical view

...of reflection seismic imaging, as practiced in the petroleum industry:

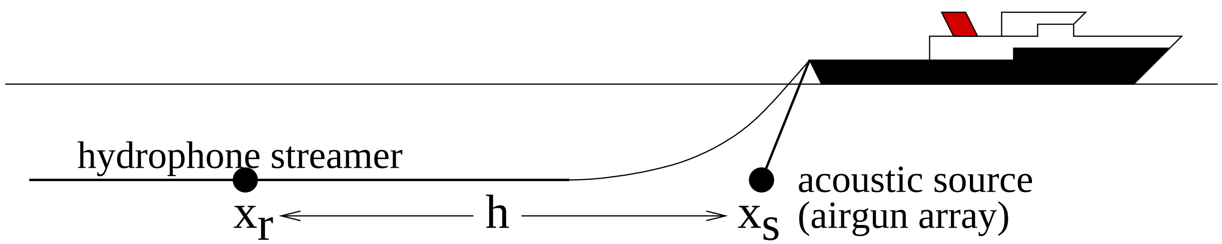
- an inverse problem, based on a model of seismic wave propagation
- contemporary practice relies on *partial linearization* and high-frequency asymptotics
- recent progress in understanding capabilities, limitations of methods based on linearization/asymptotics in presence of *strong refraction*: applications of *microlocal analysis* with implications for practice
- limitations of linearization lead to many open problems

Agenda

1. The reflection seismic experiment, nature of data and of Earth mechanical fields, the acoustic model, linearization and its limitations, definition of imaging based on high frequency asymptotics, geometric optics analysis of the model-data relationship and the GRT representation, zero-offset migration, standard processing = layered imaging
2. Analysis of GRT migration, asymptotic inversion, difficulties due to multipathing, global theory of imaging, ”wave equation” imaging;
3. The partially linearized inverse problem (“velocity analysis”), extended models, importance of invertibility, geometric optics of extensions, some invertible extensions, automating the solution of the partially linearized inverse problem via differential semblance.

Marine reflection seismology

- acoustic source (airgun array, explosives,...)
- acoustic receivers (hydrophone streamer, ocean bottom cable,...)
- recording and onboard processing



Land acquisition similar, but acquisition and processing are more complex. Vast bulk (90%+) of data acquired each year is marine.

Data parameters: time t, source location $\mathbf{x}_s$, and receiver location $\mathbf{x}_r$ or *half offset* $\mathbf{h} = \frac{\mathbf{x}_r - \mathbf{x}_s}{2}$, $h = |\mathbf{h}|$.

Idealized marine "streamer" geometry: $\mathbf{x}_s$ and $\mathbf{x}_r$ lie roughly on constant depth plane, source-receiver lines are parallel $\rightarrow$ 3 spatial degrees of freedom (eg. $\mathbf{x}_s, h$): *codimension 1*. [Other geometries are interesting, eg. ocean bottom cables, but streamer surveys still prevalent.]

How much data? Contemporary surveys may feature

- Simultaneous recording by multiple streamers (up to 12!)
- Many (roughly) parallel ship tracks ("lines"), areal coverage
- single line ("2D") $\sim$ Gbyte; multiple lines ("3D") $\sim$ Tbyte

NB: *In these lectures, will largely ignore sampling issues and treat data as continuously sampled.* First of many approximations...

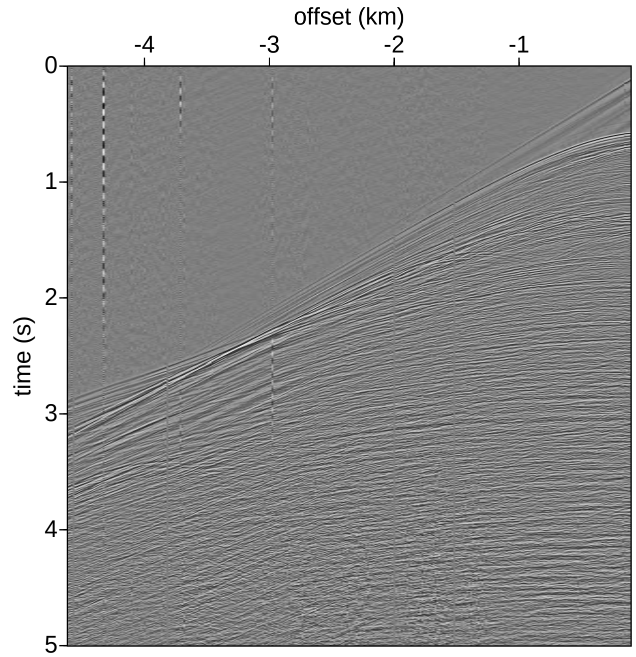
Gathers: distinguished data subsets

Aka “bins”, extracted from data after acquisition.

Characterized by common value of an acquisition parameter

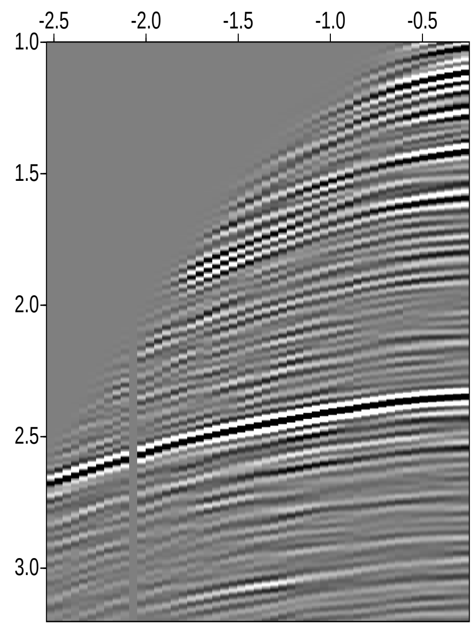
- shot (or common source) gather: traces with same shot location $\mathbf{x}_s$ (previous expls)
- offset (or common offset) gather: traces with same half offset h
- ...

Shot gather, Mississippi Canyon



(thanks: Exxon)

Lightly processed...see the waves!



bandpass filter 4-10-25-40 Hz, mute

A key observation

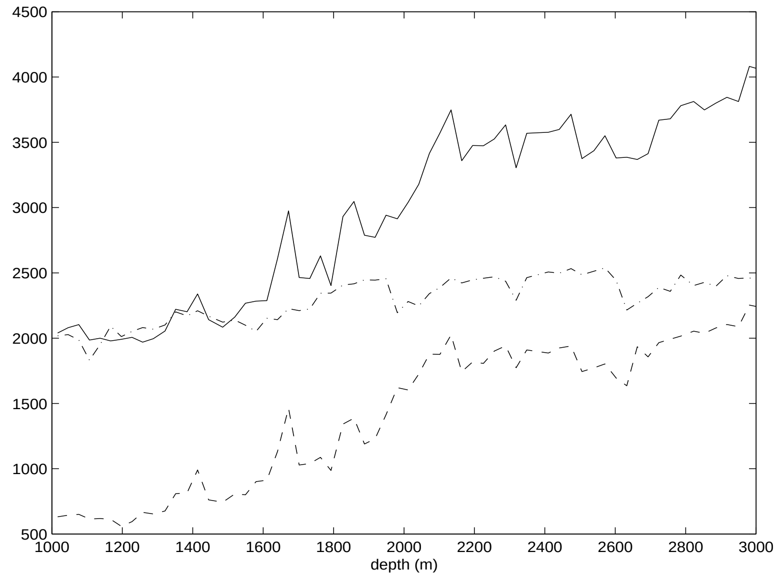
The most striking visual characteristic of seismic reflection data: presence of wave events (“reflections”) = coherent space-time structures.

What features in the subsurface structure cause reflections to occur?

Abrupt (wavelength scale) changes in material mechanics act as internal boundaries, causing reflection of waves.

What is the mechanism through which this occurs?

Well logs: a “direct” view of the subsurface



Blocked logs from well in North Sea (thanks: Mobil R & D). Solid: p-wave velocity (m/s), dashed: s-wave velocity (m/s), dash-dot: density (kg/m^3). “Blocked” means “averaged” (over 30 m windows). Original sample rate of log tool < 1 m. **Reflectors** = jumps in velocities, density, **velocity trends**.

The Modeling Task

A useful model of the reflection seismology experiment must

- predict wave motion
- produce reflections from reflectors
- accomodate significant variation of wave velocity, material density,...

A *really good* model will also accomodate

- multiple wave modes, speeds
- material anisotropy
- attenuation, frequency dispersion of waves
- complex source, receiver characteristics

The Acoustic Model

Not *really good*, but good enough for this week and basis of most contemporary processing.

Relates $\rho(\mathbf{x})$= material density, $\lambda(\mathbf{x})$ = bulk modulus, $p(\mathbf{x},t)$= pressure, $\mathbf{v}(\mathbf{x},t)$ = particle velocity, $\mathbf{f}(\mathbf{x},t)$= force density (sound source):

$$\rho\frac{\partial \mathbf{v}}{\partial t} = -\nabla p + \mathbf{f},$$

$$\frac{\partial p}{\partial t} = -\lambda \nabla \cdot \mathbf{v} \;\; (+\mathrm{i.c.}'\mathrm{s}, \mathrm{b.c.}'\mathrm{s})$$

(compressional) wave speed $c = \sqrt{\frac{\lambda}{\rho}}$

acoustic field potential $u(\mathbf{x},t) = \int_{-\infty}^{t} ds\, p(\mathbf{x},s)$:

$$p = \frac{\partial u}{\partial t},\;\; \mathbf{v} = \frac{1}{\rho}\nabla u$$

Equivalent form: second order wave equation for potential

$$\frac{1}{\rho c^2}\frac{\partial^2 u}{\partial t^2} - \nabla \cdot \frac{1}{\rho}\nabla u = \int_{-\infty}^{t} dt\,\nabla \cdot \left(\frac{\mathbf{f}}{\rho}\right) \equiv \frac{f}{\rho}$$

plus initial, boundary conditions.

Theory

Weak solution of Dirichlet problem in $\Omega \subset \mathbf{R}^3$ (similar treatment for other b. c.'s):

$$u \in C^1([0,T]; L^2(\Omega)) \cap C^0([0,T]; H^1_0(\Omega))$$

satisfying for any $\phi \in C_0^\infty((0,T) \times \Omega)$,

$$\int_0^T \int_\Omega dt\, dx \left\{ \frac{1}{\rho c^2} \frac{\partial u}{\partial t} \frac{\partial \phi}{\partial t} - \frac{1}{\rho} \nabla u \cdot \nabla \phi + \frac{1}{\rho} f \phi \right\} = 0$$

Theorem (Lions, 1972) Suppose that $\log \rho, \log c \in L^\infty(\Omega), f \in L^2(\Omega \times \mathbf{R})$. Then weak solutions of Dirichlet problem exist, uniquely determined by initial data

$$u(\cdot, 0) \in H^1_0(\Omega),\ \frac{\partial u}{\partial t}(\cdot, 0) \in L^2(\Omega)$$

NB: No hint of waves here...

Further idealizations

- density is constant,
- source force density is *isotropic point radiator with known time dependence* ("source pulse" $w(t)$)

$$f(\mathbf{x}, t; \mathbf{x}_s) = w(t)\delta(\mathbf{x} - \mathbf{x}_s)$$

$\Rightarrow$ acoustic potential, pressure depends on $\mathbf{x}_s$ also.

Forward map $\mathcal{F}[c]$ = time history of pressure for each $\mathbf{x}_s$ at receiver locations $\mathbf{x}_r$ (predicted seismic data), as function of velocity field $c(\mathbf{x})$:

$$\mathcal{F}[c] = \{p(\mathbf{x}_r, t; \mathbf{x}_s)\}$$

Reflection seismic inverse problem

given *observed seismic data* d, find c so that

$$\mathcal{F}[c] \simeq d$$

This inverse problem is

- large scale - up to Tbytes, Pflops
- nonlinear
- yields to no known direct attack

Partial linearization

Almost all useful technology to date relies on partial linearization: write $c = v(1+r)$ and treat r as relative first order perturbation about v, resulting in perturbation of presure field $\delta p = \frac{\partial \delta u}{\partial t} = 0, t \leq 0$, where

$$\left(\frac{1}{v^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\delta u = \frac{2r}{v^2}\frac{\partial^2 u}{\partial t^2}$$

Define **linearized forward map** F by

$$F[v]r = \{\delta p(\mathbf{x}_r, t; \mathbf{x}_s)\}$$

Analysis of $F[v]$ is the main content of contemporary reflection seismic theory.

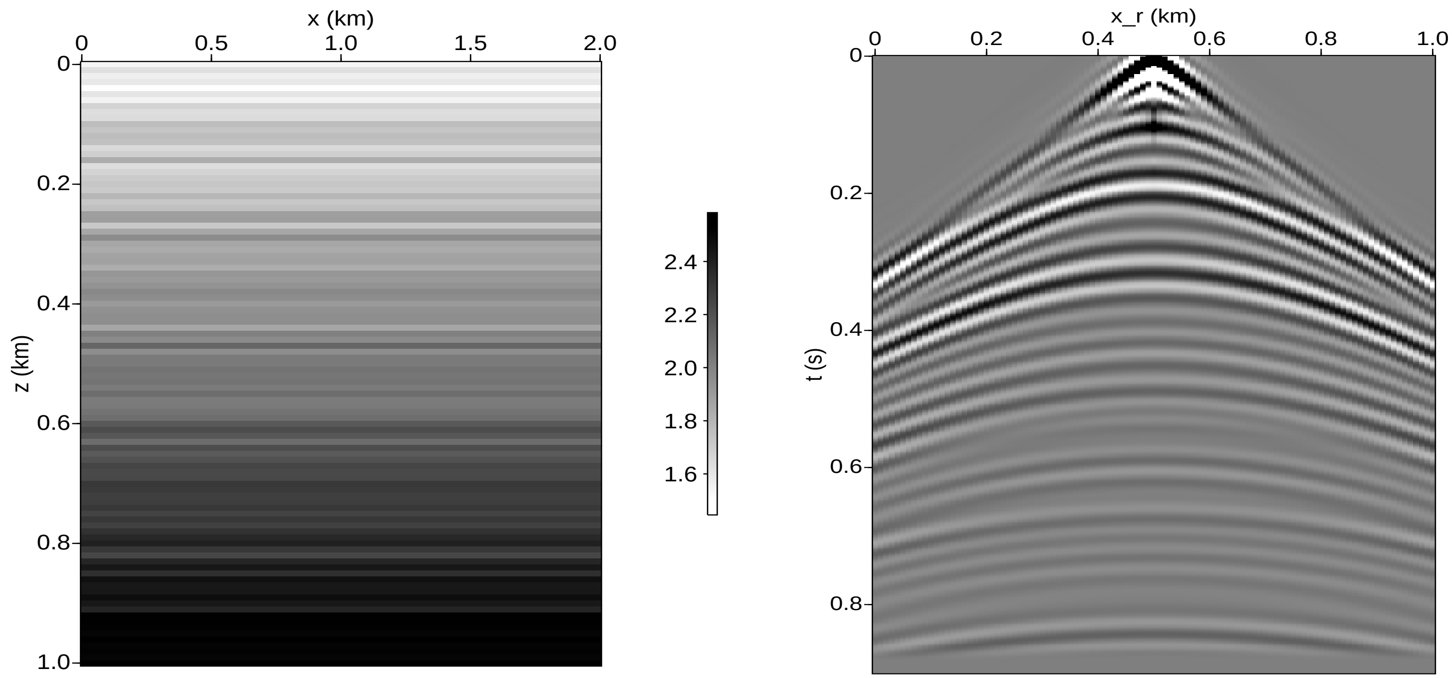
Linearization error

Critical question: If there is any justice $F[v]r$ = directional derivative $D\mathcal{F}[v][vr]$ of $\mathcal{F}$ - but in what sense? Physical intuition, numerical simulation, and not nearly enough mathematics: linearization error

$$\mathcal{F}[v(1+r)] - (\mathcal{F}[v] + F[v]r)$$

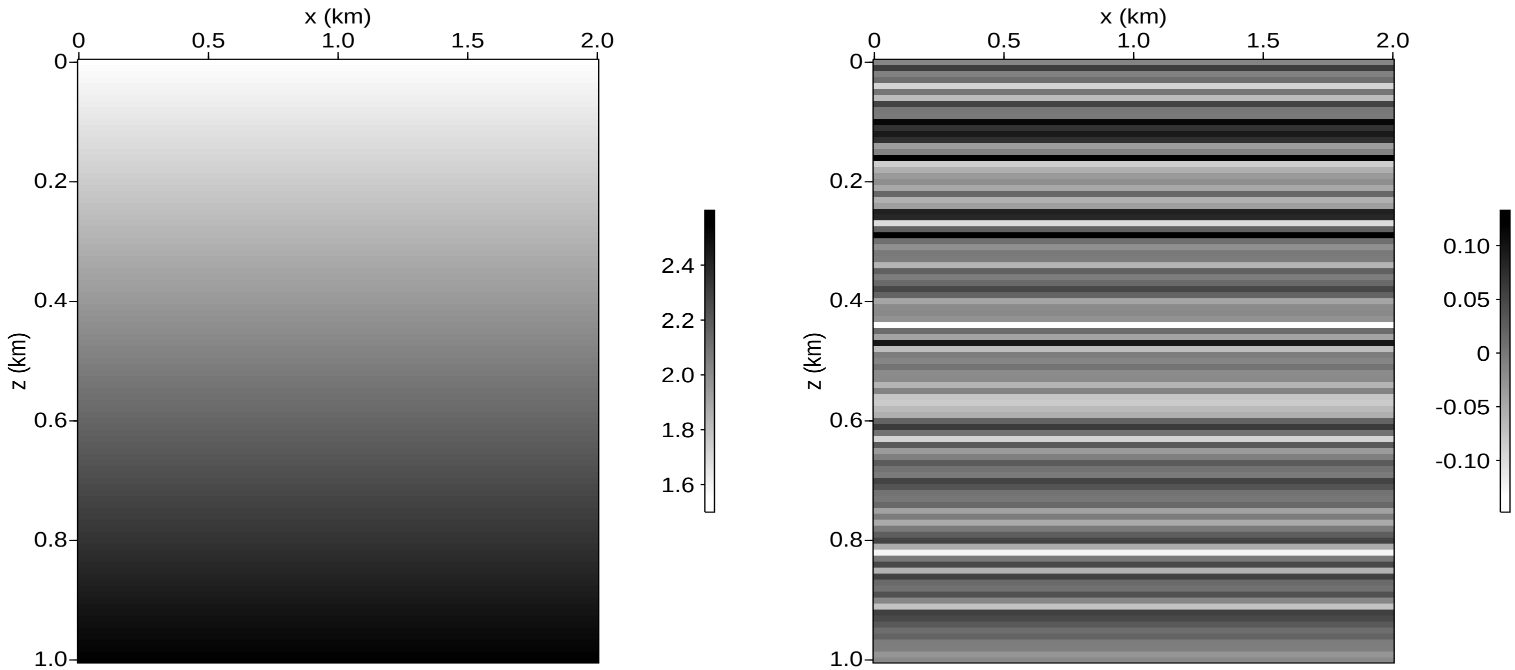
- *small* when v smooth, r rough or oscillatory on wavelength scale - well-separated scales
- *large* when v not smooth and/or r not oscillatory - poorly separated scales

2D finite difference simulation: shot gathers with typical marine seismic geometry. Smooth (linear) $v(x,z)$, oscillatory (random) $r(x,z)$ depending only on z("layered medium"). Source wavelet $w(t)$ = bandpass filter.



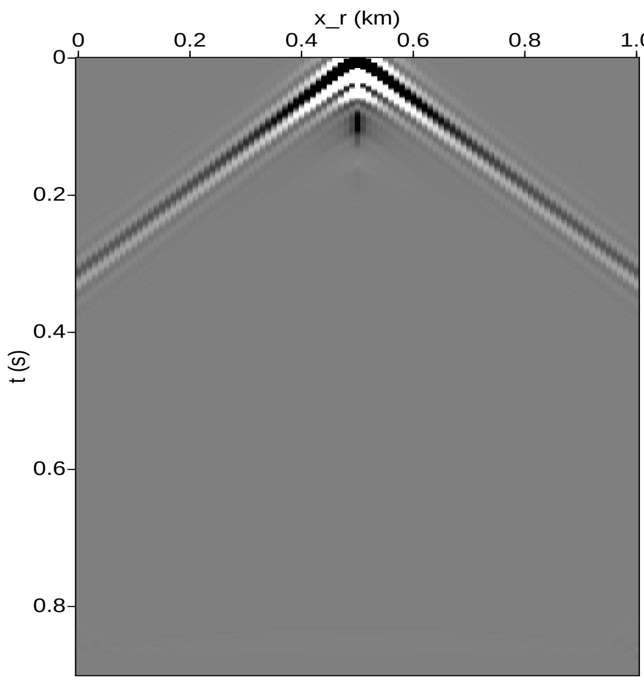
Left: Total velocity $c = v(1+r)$ with smooth (linear) background $v(x,z)$, oscillatory (random) $r(x,z)$. Std dev of r = 5%.

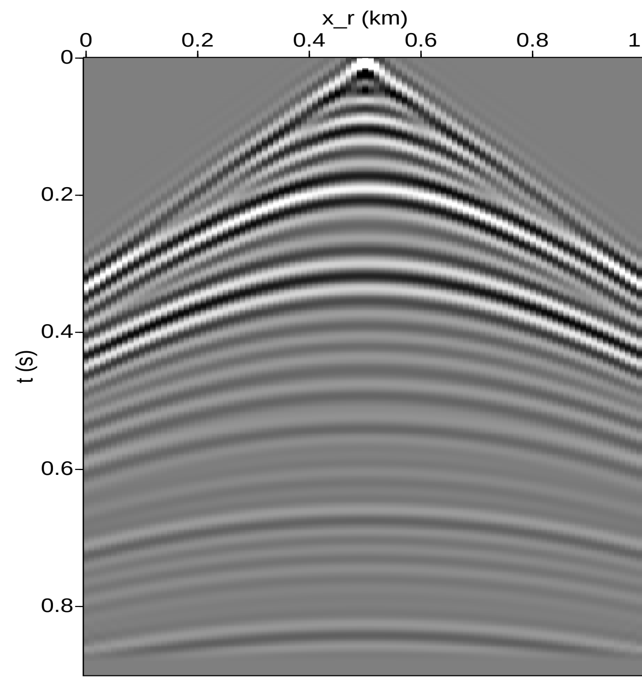
Right: Simulated seismic response ($\mathcal{F}[v(1+r)]$), wavelet = bandpass filter 4-10-30-45 Hz. Simulator is (2,4) finite difference scheme.



,

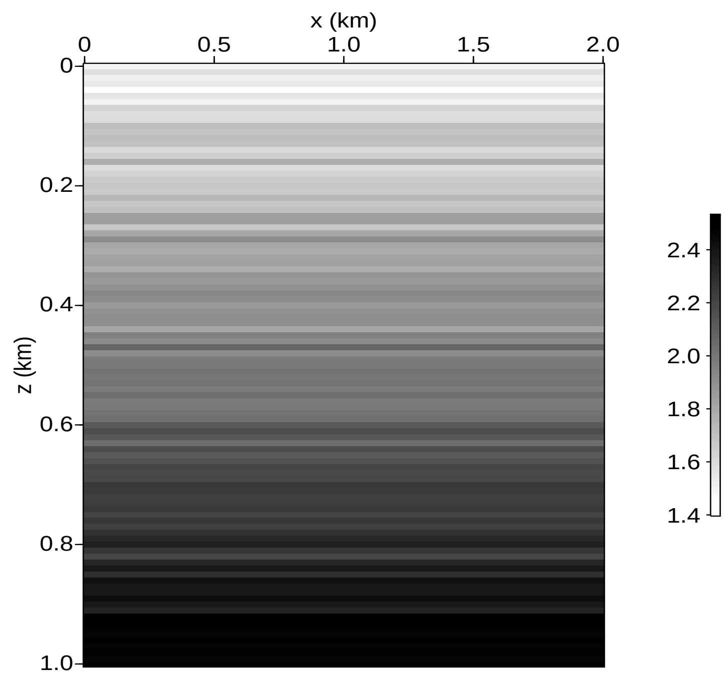
Model in previous slide as smooth background (left, $v(x, z)$) plus rough perturbation (right, $r(x, z)$).





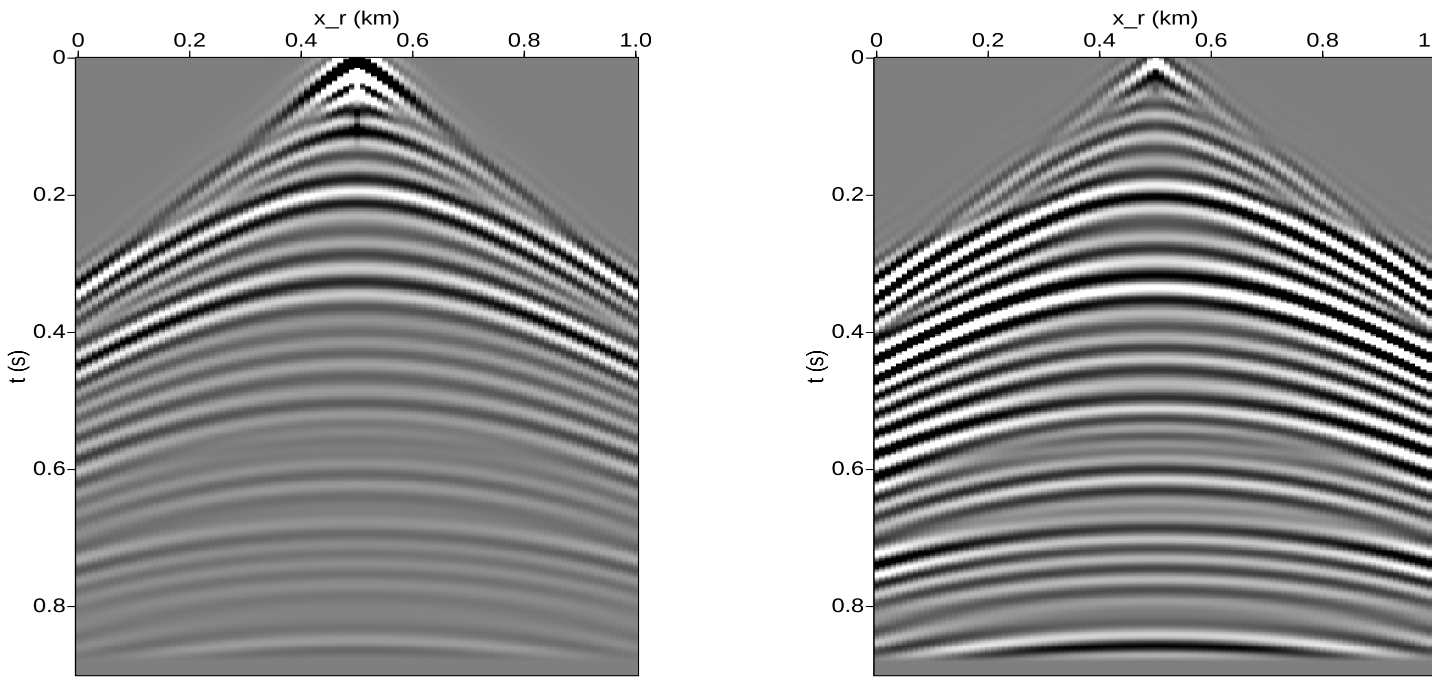
.

Left: Simulated seismic response of smooth model ($\mathcal{F}[v]$),
Right: Simulated linearized response, rough perturbation of smooth model ($F[v]r$)



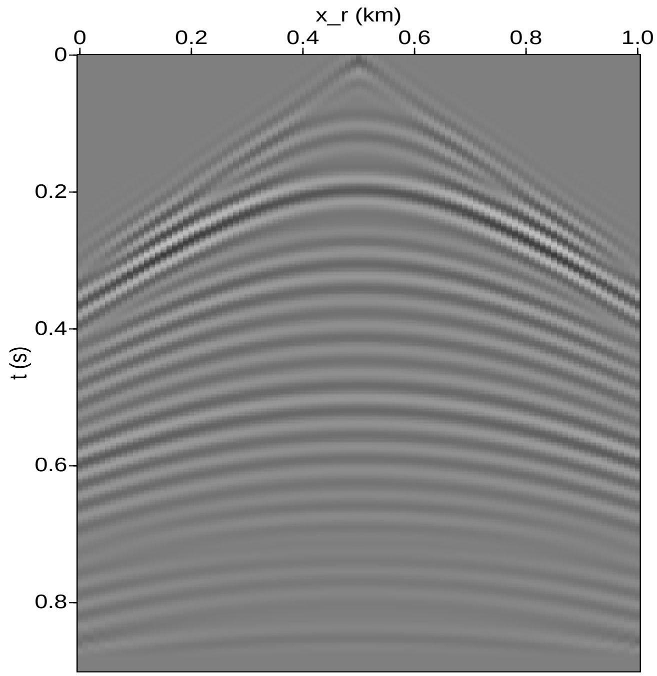
,

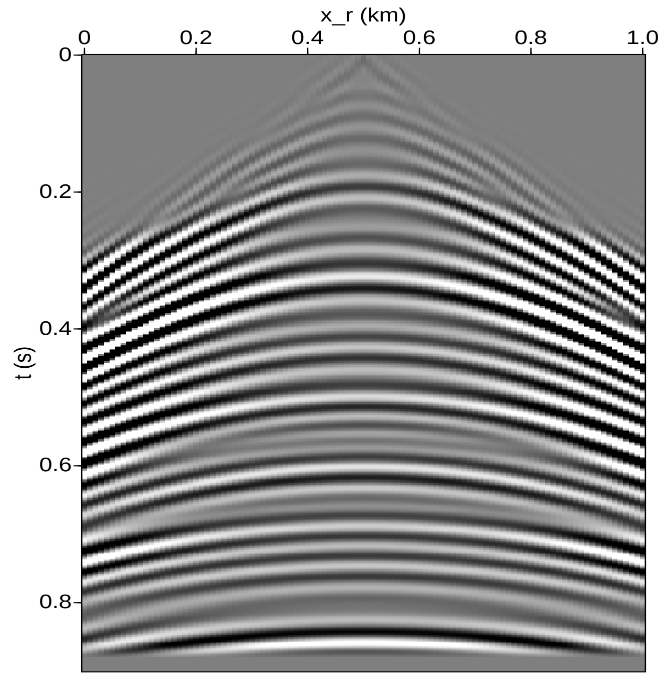
Model in previous slide as rough background (left, $v(x, z)$) plus smooth 5% perturbation ($r(x, z)$).



.

Left: Simulated seismic response of rough model ($\mathcal{F}[v]$),
Right: Simulated linearized response, smooth perturbation of rough model ($F[v]r$)





Left: linearization error ($\mathcal{F}[v(1+r)] - \mathcal{F}[v] - F[v]r$), rough perturbation of smooth background

Right: linearization error, smooth perturbation of rough background (plotted with same grey scale).

Summary

- v smooth, r oscillatory $\Rightarrow$ $F[v]r$ approximates **primary reflection** = result of wave interacting with material heterogeneity only once (single scattering); error consists of **multiple reflections**, which are "not too large" if r is "not too big", and sometimes can be suppressed.
- v nonsmooth, r smooth $\Rightarrow$ error consists of *time shifts* in waves which are very large perturbations as waves are oscillatory.

No mathematical results are known which justify/explain these observations in any rigorous way, except in 1D.

Velocity Analysis and Imaging

Velocity analysis problem = partially linearized inverse problem: given d find v, r so that

$$S[v] + F[v]r \simeq d$$

Imaging problem = linear subproblem: given d and v, find r so that

$$F[v]r \simeq d - S[v]$$

Last 20 years:

- much progress on imaging
- much less on velocity analysis

Aymptotic assumption

Linearization is accurate $\Leftrightarrow$ length scale of $v >>$ length scale of $r \simeq$ wavelength, properties of $F[v]$ dominated by those of $F_\delta[v]$ (= $F[v]$ with $w = \delta$). Implicit in migration concept (eg. Hagedoorn, 1954); explicit use: Cohen & Bleistein, SIAM JAM 1977.

Key idea: **reflectors** (rapid changes in r) emulate *singularities*; **reflections** (rapidly oscillating features in data) also emulate singularities.

NB: "everybody's favorite reflector": the smooth interface across which r jumps. *But* this is an oversimplification - reflectors in the Earth may be complex zones of rapid change, pehaps in all directions. More flexible notion needed!!

Wave Front Sets

Recall characterization of smoothness via Fourier transform: $u \in \mathcal{D}'(\mathbf{R}^n)$ is smooth at $\mathbf{x}_0 \Leftrightarrow$ for some nbhd X of $\mathbf{x}_0$, any $\phi \in \mathcal{E}(X)$ and N, there is $C_N \geq 0$ so that for any $\boldsymbol{\xi} \neq 0$,

$$\left|\mathcal{F}(\phi u)\left(\tau\frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}\right)\right| \leq C_N \tau^{-N}$$

Harmonic analysis of singularities, *après* Hörmander: the **wave front set** $WF(u) \subset \mathbf{R}^n \times \mathbf{R}^n - \{0\}$ of $u \in \mathcal{D}'(\mathbf{R}^n)$ - captures orientation as well as position of singularities.

$(\mathbf{x}_0, \boldsymbol{\xi}_0) \notin WF(u) \Leftrightarrow$, there is some open nbhd $X \times \Xi \subset \mathbf{R}^n \times \mathbf{R}^n - \{0\}$ of $(\mathbf{x}_0, \xi_0)$ so that for any $\phi \in \mathcal{E}(X)$, N, there is $C_N \geq 0$ so that for all $\boldsymbol{\xi} \in \Xi$,

$$\left|\mathcal{F}(\phi u)\left(\tau\frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}\right)\right| \leq C_N \tau^{-N}$$

Housekeeping chores

(i) note that the nbhds Ξ may naturally be taken to be *cones*;

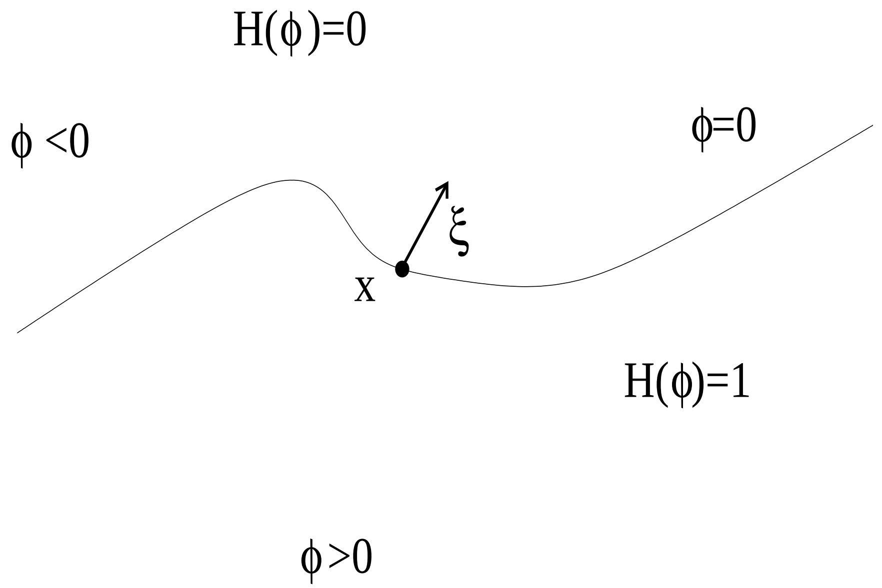
(ii) u is smooth at $\mathbf{x}_0 \Leftrightarrow (\mathbf{x}_0, \boldsymbol{\xi}_0) \notin WF(u)$ for all $\boldsymbol{\xi}_0 \in \mathbf{R}^n - \{0\}$;

(iii) $WF(u)$ is invariant under chg. of coords if it is regarded as a subset of the *cotangent bundle* $T^*(\mathbf{R}^n)$ (i.e. the ξ components transform as covectors).

[Good refs: Duistermaat, 1996; Taylor, 1981; Hörmander, 1983]

The standard example: if u jumps across the interface $f(\mathbf{x}) = 0$, otherwise smooth, then $WF(u) \subset \mathcal{N}_f = \{(\mathbf{x}, \boldsymbol{\xi}) : f(\mathbf{x}) = 0,\ \boldsymbol{\xi} || \nabla f(\mathbf{x})\}$ (*normal bundle* of $f = 0$).

Wavefront set of a jump discontinuity



$$WF(H(\phi)) = \{(\mathbf{x}, \boldsymbol{\xi}) : \phi(\mathbf{x}) = 0,\ \boldsymbol{\xi} || \nabla\phi(\mathbf{x})\}$$

Microlocal property of differential operators

Suppose $u \in \mathcal{D}'(\mathbf{R}^n)$, $(\mathbf{x}_0, \boldsymbol{\xi}_0) \notin WF(u)$, and $P(\mathbf{x}, D)$ is a partial differential operator:

$$P(\mathbf{x}, D) = \sum_{|\alpha| \leq m} a_\alpha(x) D^\alpha$$

$$D = (D_1, ..., D_n), \;\; D_i = -i \frac{\partial}{\partial x_i}$$

$$\alpha = (\alpha_1, ..., \alpha_n), \;\; |\alpha| = \sum_i \alpha_i,$$

$$D^\alpha = D_1^{\alpha_1} ... D_n^{\alpha_n}$$

Then $(\mathbf{x}_0, \xi_0) \notin WF(P(\mathbf{x}, D)u)$ [i.e.: $WF(Pu) \subset WF(u)$].

Proof

Choose $X \times \Xi$ as in the definition, $\phi \in \mathcal{D}(X)$ form the required Fourier transform

$$\int dx\, e^{i\mathbf{x}\cdot(\tau\xi)}\phi(\mathbf{x})P(\mathbf{x}, D)u(\mathbf{x})$$

and start integrating by parts: eventually

$$= \sum_{|\alpha|\leq m} \tau^{|\alpha|}\xi^{\alpha} \int dx\, e^{i\mathbf{x}\cdot(\tau\xi)}\phi_{\alpha}(\mathbf{x})u(\mathbf{x})$$

where $\phi_{\alpha} \in \mathcal{D}(X)$ is a linear combination of derivatives of ϕ and the a_{α}s. Since each integral is rapidly decreasing as $\tau \to \infty$ for $\xi \in \Xi$, it remains rapidly decreasing after multiplication by $\tau^{|\alpha|}$, and so does the sum. **Q. E. D.**

Formalizing the reflector concept

Key idea, restated: reflectors (or "reflecting elements") will be points in $WF(r)$. Reflections will be points in $WF(d)$.

These ideas lead to a usable definition of *image*: a reflectivity model $\tilde{r}$ is an image of r if $WF(\tilde{r}) \subset WF(r)$ (the closer to equality, the better the image).

Idealized **migration problem**: given d (hence $WF(d)$) deduce somehow a function which has *the right reflectors*, i.e. a function $\tilde{r}$ with $WF(\tilde{r}) \simeq WF(r)$.

NB: you're going to need v! ("It all depends on v(x,y,z)" - J. Claerbout)

Integral representation of linearized operator

With $w = \delta$, acoustic potential u is same as Causal Green's function $G(\mathbf{x}, t; \mathbf{x}_s)$ = retarded fundamental solution:

$$\left(\frac{1}{v^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) G(\mathbf{x}, t; \mathbf{x}_s) = \delta(t)\delta(\mathbf{x} - bx_s)$$

and $G \equiv 0, t < 0$. Then ($w = \delta$!) $p = \frac{\partial G}{\partial t}$, $\delta p = \frac{\partial \delta G}{\partial t}$, and

$$\left(\frac{1}{v^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) \delta G(\mathbf{x}, t; \mathbf{x}_s) = \frac{2}{v^2(\mathbf{x})}\frac{\partial^2 G}{\partial t^2}(\mathbf{x}, t; \mathbf{x}_s) r(\mathbf{x})$$

Simplification: from now on, define $F[v]r = \delta G|_{\mathbf{x}=\mathbf{x}_r}$ - i.e. lose a t-derivative. Duhamel's principle $\Rightarrow$

$$\delta G(\mathbf{x}_r, t; \mathbf{x}_s) = \int dx \, \frac{2r(\mathbf{x})}{v(\mathbf{x})^2} \int ds \, G(\mathbf{x}_r, t - s; \mathbf{x}) \frac{\partial^2 G}{\partial t^2}(\mathbf{x}, s; \mathbf{x}_s)$$

Add geometric optics...

Geometric optics approximation of G should be good, as v is smooth. Summary: if $\mathbf{x}$ "not too far" from $\mathbf{x}_s$, then

$$G(\mathbf{x}, t; \mathbf{x}_s) = a(\mathbf{x}; \mathbf{x}_s)\delta(t - \tau(\mathbf{x}; \mathbf{x}_s)) + R(\mathbf{x}, t; \mathbf{x}_s)$$

where the traveltime $\tau(\mathbf{x}; \mathbf{x}_s)$ solves the eikonal equation

$$v|\nabla\tau| = 1, \;\; \tau(\mathbf{x}; \mathbf{x}_s) \sim \frac{|\mathbf{x} - \mathbf{x}_s|}{v(\mathbf{x}_s)}, \;\; \mathbf{x} \rightarrow \mathbf{x}_s$$

and the amplitude $a(\mathbf{x}; \mathbf{x}_s)$ solves the transport equation

$$\nabla \cdot (a^2 \nabla\tau) = 0, \ldots$$

Refs: Courant & Hilbert, Friedlander *Sound Pulses*, WWS *Foundations* and many refs cited there...

Simple Geometric Optics

"Not too far" means: there should be one and only one ray of geometric optics connecting each $\mathbf{x}_s$ or $\mathbf{x}_r$ to each $\mathbf{x} \in \mathrm{supp} r$.

Will call this the **simple geometric optics** assumption.

Within region satisfying simple geometric optics assumption, τ is smooth ($\mathbf{x} \neq \mathbf{x}_s$) solution of eikonal equation. Effective methods for numerical solution of eikonal, transport equations: ray tracing (Lagrangian), various sorts of upwind finite difference (Eulerian) methods. See eg. Sethian book, WWS 1999 MGSS notes (online) for details.

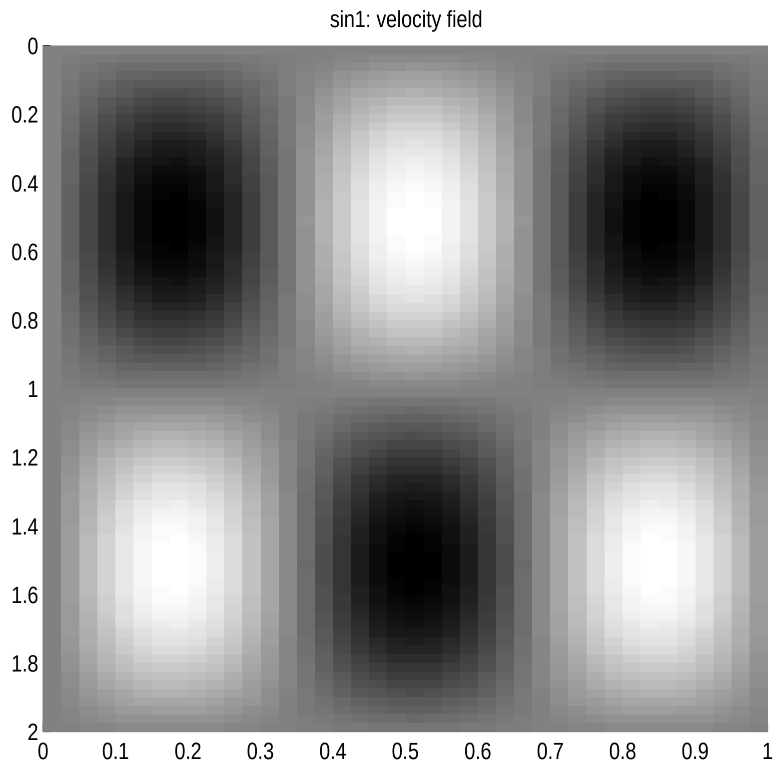
Caution - caustics!

For "random but smooth" $v(\mathbf{x})$ with variance σ, more than one connecting ray occurs as soon as the distance is $O(\sigma^{-2/3})$. Such *multipathing* is invariably accompanied by the formation of a *caustic* = envelope of rays (White, 1982).

Upon caustic formation, the simple geometric optics field description above is no longer correct.

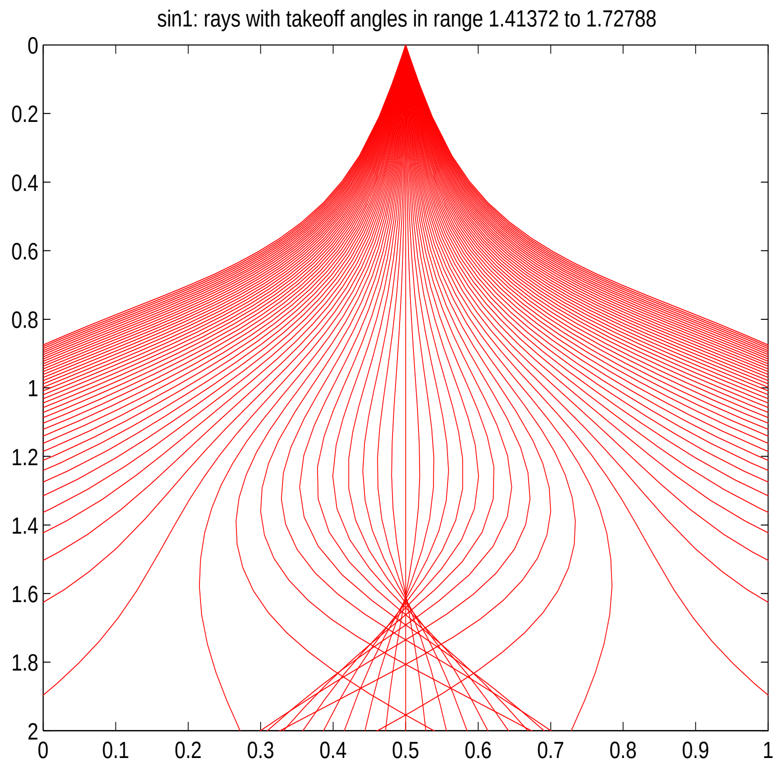
Failure of GO at caustic understood in 19th century. Generalization of GO to regions containing caustics accomplished by Ludwig and Kravtsov, 1966-7, elaborated by Maslov, Hörmander, Duistermaat, many others.

A caustic example (1)



2D Example of strong refraction: Sinusoidal velocity field $v(x, z) = 1{+}0.2 \sin \frac{\pi z}{2} \sin 3\pi x$

A caustic example (2)



Rays in sinusoidal velocity field, source point = origin. Note formation of caustic, multiple rays to source point in lower center.

An oft-forgotten detail

All of this is meaningful only if the remainder R is small in a suitable sense: energy estimate (**Exercise!**) $\Rightarrow$

$$\int dx \int_0^T dt\, |R(\mathbf{x},t;\mathbf{x}_s)|^2 \le C\|v\|_{\mathrm{C}^4}$$

(this is an easy, suboptimal estimate - with more work can replace 4 with 2)

If $v \in C^\infty$, can complete the geometric optics approximation of the Green's function so that the difference is C^∞ - then the two sides have the same singularities, ie. the same wavefront set.

Finally, a wave!

The geometric optics approximation to the Green's function

$$G(\mathbf{x}, t; \mathbf{x}_s) \simeq a(\mathbf{x}; \mathbf{x}_s)\delta(t - \tau(\mathbf{x}; \mathbf{x}_s))$$

describes a (singular) quasi-spherical waves [spherical, if v = const., for then $\tau(\mathbf{x}, \mathbf{x}_s) = |\mathbf{x} - \mathbf{x}_s|/v$].

Geometric optics is the the best currently available explanation for waves in heterogeneous media. Note the inadequacy: v must be *smooth*, but the compressional velocity distribution in the Earth varies on all scales!

The linearized operator as Generalized Radon Transform

Assume: $\mathrm{supp}\, r$ contained in simple geometric optics domain (each point reached by unique ray from any source or receiver point).

Then distribution kernel K of $F[v]$ is

$$K(\mathbf{x}_r, t, \mathbf{x}_s; \mathbf{x}) = \int ds\, G(\mathbf{x}_r, t-s; \mathbf{x}) \frac{\partial^2 G}{\partial t^2}(\mathbf{x}, s; \mathbf{x}_s) \frac{2}{v^2(\mathbf{x})}$$

$$\simeq \int ds\, \frac{2a(\mathbf{x}_r, \mathbf{x})a(\mathbf{x}, \mathbf{x}_s)}{v^2(\mathbf{x})} \delta'(t - s - \tau(\mathbf{x}_r, \mathbf{x}))\delta''(s - \tau(\mathbf{x}, \mathbf{x}_s))$$

$$= \frac{2a(\mathbf{x},\mathbf{x}_r)a(\mathbf{x},\mathbf{x}_s)}{v^2(\mathbf{x})}\delta''(t-\tau(\mathbf{x},\mathbf{x}_r)-\tau(\mathbf{x},\mathbf{x}_s))$$

provided that

$$\nabla_{\mathbf{x}}\tau(\mathbf{x},\mathbf{x}_r)+\nabla_{\mathbf{x}}\tau(\mathbf{x},\mathbf{x}_s)\neq 0$$

$\Leftrightarrow$ velocity at $\mathbf{x}$ of ray from $\mathbf{x}_s$ **not** negative of velocity of ray from $\mathbf{x}_r$ $\Leftrightarrow$ *no forward scattering*. [Gel'fand and Shilov, 1958 - when is pullback of distribution again a distribution].

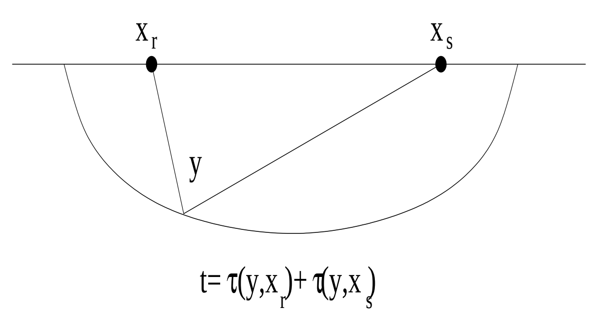
GRT = “Kirchhoff” modeling

So: for r supported in simple geometric optics domain, no forward scattering $\Rightarrow$

$$\delta G(\mathbf{x}_r, t; \mathbf{x}_s) \simeq$$

$$\frac{\partial^2}{\partial t^2} \int dx \, \frac{2r(\mathbf{x})}{v^2(\mathbf{x})} a(\mathbf{x}, \mathbf{x}_r) a(\mathbf{x}, \mathbf{x}_s) \delta(t - \tau(\mathbf{x}, \mathbf{x}_r) - \tau(\mathbf{x}, \mathbf{x}_s))$$

That is: pressure perturbation is sum (integral) of r over *reflection isochron* $\{\mathbf{x} : t = \tau(\mathbf{x}, \mathbf{x}_r) + \tau(\mathbf{x}, \mathbf{x}_s)\}$, w. weighting, filtering. Note: if v =const. then isochron is ellipsoid, as $\tau(\mathbf{x}_s, \mathbf{x}) = |\mathbf{x}_s - \mathbf{x}|/v$!



Zero Offset data and the Exploding Reflector

Zero offset data ($\mathbf{x}_s = \mathbf{x}_r$) is seldom actually measured (contrast radar, sonar!), but routinely *approximated* through *NMO-stack* (to be explained later).

Extracting image from zero offset data, rather than from all (100's) of offsets, is tremendous *data reduction* - when approximation is accurate, leads to excellent images.

Imaging basis: the *exploding reflector* model (Claerbout, 1970's).

For zero-offset data, distribution kernel of $F[v]$ is

$$K(\mathbf{x}_s, t, \mathbf{x}_s; \mathbf{x}) = \frac{\partial^2}{\partial t^2} \int \, ds \, \frac{2}{v^2(\mathbf{x})} G(\mathbf{x}_s, t-s; \mathbf{x}) G(\mathbf{x}, s; \mathbf{x}_s)$$

Under some circumstances (explained below), K (= G time-convolved with itself) is "similar" (also explained) to $\tilde{G}$ = Green's function for $v/2$. Then

$$\delta G(\mathbf{x}_s, t; \mathbf{x}_s) \sim \frac{\partial^2}{\partial t^2} \int \, dx \, \tilde{G}(\mathbf{x}_s, t, \mathbf{x}) \frac{2r(\mathbf{x})}{v^2(\mathbf{x})}$$

$\sim$ solution w of

$$\left(\frac{4}{v^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) w = \delta(t) \frac{2r}{v^2}$$

Thus reflector "explodes" at time zero, resulting field propagates in "material" with velocity $v/2$.

Explain when the exploding reflector model "works", i.e. when G time-convolved with itself is "similar" to $\tilde{G}$ = Green's function for $v/2$. If supp r lies in simple geometry domain, then

$$K(\mathbf{x}_s, t, \mathbf{x}_s; \mathbf{x}) = \int ds \, \frac{2a^2(\mathbf{x}, \mathbf{x}_s)}{v^2(\mathbf{x})} \delta(t - s - \tau(\mathbf{x}_s, \mathbf{x}))\delta''(s - \tau(\mathbf{x}, \mathbf{x}_s))$$

$$= \frac{2a^2(\mathbf{x}, \mathbf{x}_s)}{v^2(\mathbf{x})} \delta''(t - 2\tau(\mathbf{x}, \mathbf{x}_s))$$

whereas the Green's function $\tilde{G}$ for $v/2$ is

$$\tilde{G}(\mathbf{x}, t; \mathbf{x}_s) = \tilde{a}(\mathbf{x}, \mathbf{x}_s)\delta(t - 2\tau(\mathbf{x}, \mathbf{x}_s))$$

(half velocity = double traveltime, same rays!).

Difference between effects of K, $\tilde{G}$: for each $\mathbf{x}_s$ scale r by smooth fcn - preserves $WF(r)$ hence $WF(F[v]r)$ and relation between them. Also: adjoints have same effect on WF sets.

Upshot: from imaging point of view (i.e. apart from amplitude, derivative (filter)), kernel of $F[v]$ restricted to zero offset is same as Green's function for $v/2$, *provided that simple geometry hypothesis holds:* only one ray connects each source point to each scattering point, ie. *no multipathing*.

See Claerbout, IEI, for examples which demonstrate that multipathing really does invalidate exploding reflector model.

Standard Processing

Inspirational interlude: the sort-of-layered theory =“Standard Processing”

Suppose were v,r functions of $z = x_3$ only, all sources and receivers at $z = 0$. Then the entire system is translation-invariant in $x_1, x_2 \Rightarrow$ Green’s function G its perturbation δG, and the idealized data $\delta G|_{z=0}$ are really only functions of t, z, and *half-offset* $h = |\mathbf{x}_s - \mathbf{x}_r|/2$. There would be *only one seismic experiment*, equivalent to any *common midpoint gather* (“CMP”).

This isn’t really true - *look at the data!!!* However it is *approximately* correct in many places in the world: CMPs change very slowly with midpoint $\mathbf{x}_m = (\mathbf{x}_r + \mathbf{x}_s)/2$.

Standard processing: treat each CMP *as if it were the result of an experiment performed over a layered medium*, but permit the layers to vary with midpoint.

Thus $v = v(z), r = r(z)$ for purposes of analysis, but at the end $v = v(\mathbf{x}_m, z), r = r(\mathbf{x}_m, z)$.

$$F[v]r(\mathbf{x}_r, t; \mathbf{x}_s)$$

$$\simeq \int dx \frac{2r(z)}{v^2(z)} a(\mathbf{x}, x_r) a(\mathbf{x}, x_s) \delta''(t - \tau(\mathbf{x}, x_r) - \tau(\mathbf{x}, x_s))$$

$$= \int dz \, \frac{2r(z)}{v^2(z)} \int d\omega \int dx \omega^2 a(\mathbf{x}, x_r) a(\mathbf{x}, x_s) e^{i\omega(t - \tau(\mathbf{x}, x_r) - \tau(\mathbf{x}, x_s))}$$

Since we have already thrown away smoother (lower frequency) terms, do it again using *stationary phase.* Upshot (see 2000 MGSS notes for details): up to smoother (lower frequency) error,

$$F[v]r(h,t) \simeq A(z(h,t),h)R(z(h,t))$$

Here $z(h,t)$ is the inverse of the 2-way traveltime

$$t(h,z) = 2\tau((h,0,z),(0,0,0))$$

i.e. $z(t(h,z'),h) = z'$. R is (yet another version of) “reflectivity”

$$R(z) = \frac{1}{2}\frac{dr}{dz}(z)$$

That is, $F[v]$ *is a a derivative followed by a change of variable followed by multiplication by a smooth function*. Substitute t_0 (vertical travel time) for z (depth) and you get “Inverse NMO” ($t_0 \rightarrow (t,h)$). Will be sloppy and call $z \rightarrow (t,h)$ INMO.

Anatomy of an adjoint

$$\int dt \int dh\, d(t,h) F[v] r(t,h) = \int dt \int dh\, d(t,h) A(z(t,h),h) R(z(t,h))$$

$$= \int dz\, R(z) \int dh\, \frac{\partial t}{\partial z}(z,h) A(z,h) d(t(z,h),h) = \int dz\, r(z) (F[v]^* d)(z)$$

so $F[v]^* = -\frac{\partial}{\partial z} S M[v] N[v]$, where

- $N[v]$ = **NMO operator** $N[v]d(z,h) = d(t(z,h),h)$
- $M[v]$ = multiplication by $\frac{\partial t}{\partial z} A$
- S = **stacking operator** $Sf(z) = \int dh\, f(z,h)$

Normal Op is PDO $\Rightarrow$ Imaging

$$F[v]^*F[v]r(z) = -\frac{\partial}{\partial z}\left[\int dh\,\frac{dt}{dz}(z,h)A^2(z,h)\right]\frac{\partial}{\partial z}r(z)$$

Microlocal property of PDOs $\Rightarrow WF(F[v]^*F[v]r) \subset WF(r)$ i.e. $F[v]^*$ *is an imaging operator*.

If you leave out the amplitude factor ($M[v]$) and the derivatives, as is commonly done, then you get essentially the same expression - so (NMO, stack) is an imaging operator!

It's even easy to get an (asymptotic) inverse out of this - exercise for the reader.

Now make everything dependent on $\mathbf{x}_m$ and you've got standard processing. (end of layered interlude).

But the Earth is not layered!

In general,

Is $F[v]^*$ an imaging operator?

What sort of thing is $F[v]^*F[v]$??

Stay tuned!

Mathematics of Seismic Imaging
Part II

William W. Symes

PIMS, June 2005

Review: Normal Operators and imaging

If $d = F[v]r$, then

$$F[v]^* d = F[v]^* F[v] r$$

Recall: In the layered case, $F[v]^* F[v]$ is an operator which preserves wave front sets.

Whenever $F[v]^* F[v]$ preserves wave front sets, $F[v]^*$ is an imaging operator.

Review: Generalized Radon Representation

Assume (1) r (oscillatory) supported in simple geometric optics domain for v (smooth), (2) no forward scattering. Then

$$F[v]r(\mathbf{x}_r, t; \mathbf{x}_s) \simeq$$

$$\int dx \, \frac{2r(\mathbf{x})}{v^2(\mathbf{x})} a(\mathbf{x}, \mathbf{x}_r) a(\mathbf{x}, \mathbf{x}_s) \delta''(t - \tau(\mathbf{x}, \mathbf{x}_r) - \tau(\mathbf{x}, \mathbf{x}_s))$$

Similar representation of adjoint follows:

$$F[v]^* d(\mathbf{x}) = \int\int\int dx_r \, dx_s \, dt \, a(\mathbf{x}, \mathbf{x}_r) a(\mathbf{x}, \mathbf{x}_s) \delta''(t - \tau(\mathbf{x}; \mathbf{x}_s) - \tau(\mathbf{x}; \mathbf{x}_r)) d(\mathbf{x}_r, t; \mathbf{x}_s)$$

Beylkin, *J. Math. Phys.* 1985

For r supported in simple geometric optics domain,

- $WF(F[v]^*F[v]r) \subset WF(r)$
- if $d = \mathcal{F}[v] + F[v]r$ (data consistent with linearized model), then $F[v]^*(d - \mathcal{F}[v])$ is an image of r
- an operator $F[v]^\dagger$ exists for which $F[v]^\dagger(d - \mathcal{F}[v]) - r$ is *smoother* than r, under some constraints on r - an *inverse modulo smoothing operators* or *parametrix*.

Outline of proof

Express $F[v]^*F[v]$ as "Kirchhoff modeling" followed by "Kirchhoff migration"; (ii) introduce Fourier transform; (iii) approximate for large wavenumbers using stationary phase, leads to representation of $F[v]^*F[v]$ modulo smoothing error as *pseudodifferential operator* ("ΨDO"):

$$F[v]^*F[v]r(\mathbf{x}) \simeq p(\mathbf{x}, D)r(\mathbf{x}) \equiv \int d\xi\, p(\mathbf{x}, \boldsymbol{\xi})e^{i\mathbf{x}\cdot\boldsymbol{\xi}}\hat{r}(\boldsymbol{\xi})$$

in which $p \in C^\infty$, and for some m (the *order* of p), all multiindices α, β, and all compact $K \subset \mathbf{R}^n$, there exist constants $C_{\alpha,\beta,K} \geq 0$ for which

$$|D_{\mathbf{x}}^\alpha D_{\boldsymbol{\xi}}^\beta p(\mathbf{x}, \boldsymbol{\xi})| \leq C_{\alpha,\beta,K}(1 + |\boldsymbol{\xi}|)^{m-|\beta|},\ \mathbf{x} \in K$$

Explicit computation of **symbol** p - for details, see Notes on Math Foundations.

Microlocal Propertyof ΨDOs

:

if $p(x,D)$ is a ΨDO, $u \in \mathcal{E}'(\mathbf{R}^n)$ then $WF(p(x,D)u) \subset WF(u)$.

Will prove this, from which imaging property of prestack Kirchhoff migration follows. First, a few other properties:

- differential operators are ΨDOs (easy - exercise)
- ΨDOs of order m form a module over $C^\infty(\mathbf{R}^n)$ (also easy)
- product of ΨDO order m, ΨDO order l = ΨDO order $\leq m+l$; adjoint of ΨDO order m is ΨDO order m (harder)

Complete accounts of theory, many apps: books of Duistermaat, Taylor, Nirenberg, Treves, Hörmander.

Proof of Microlocal Property

Suppose $(\mathbf{x}_0, \boldsymbol{\xi}_0) \notin WF(u)$, choose neighborhoods X, Ξ as in defn, with Ξ conic. Need to choose analogous nbhds for $P(x, D)u$. Pick $\delta > 0$ so that $B_{3\delta}(\mathbf{x}_0) \subset X$, set $X' = B_\delta(\mathbf{x}_0)$.

Similarly pick $0 < \epsilon < 1/3$ so that $B_{3\epsilon}(\boldsymbol{\xi}_0/|\boldsymbol{\xi}_0|) \subset \Xi$, and chose $\Xi' = \{\tau\boldsymbol{\xi} : \boldsymbol{\xi} \in B_\epsilon(\boldsymbol{\xi}_0/|\boldsymbol{\xi}_0|), \tau > 0\}$.

Need to choose $\phi \in \mathcal{E}'(X')$, estimate $\mathcal{F}(\phi P(\mathbf{x}, D)u)$. Choose $\psi \in \mathcal{E}(X)$ so that $\psi \equiv 1$ on $B_{2\delta}(\mathbf{x}_0)$.

NB: this implies that if $\mathbf{x} \in X'$, $\psi(\mathbf{y}) \neq 1$ then $|\mathbf{x} - \mathbf{y}| \geq \delta$.

Write $u = (1-\psi)u + \psi u$. Claim: $\phi P(\mathbf{x}, D)((1-\psi)u)$ is smooth.

$$\phi(\mathbf{x})P(\mathbf{x}, D)((1-\psi)u))(\mathbf{x})$$

$$= \phi(\mathbf{x}) \int d\xi \, P(\mathbf{x}, \boldsymbol{\xi}) e^{i\mathbf{x}\cdot\boldsymbol{\xi}} \int dy \, (1-\psi(\mathbf{y}))u(\mathbf{y})e^{-i\mathbf{y}\cdot\boldsymbol{\xi}}$$

$$= \int d\xi \int dy \, P(\mathbf{x}, \boldsymbol{\xi})\phi(\mathbf{x})(1-\psi(\mathbf{y}))e^{i(\mathbf{x}-\mathbf{y})\cdot\boldsymbol{\xi}}u(\mathbf{y})$$

$$= \int d\xi \int dy \, (-\nabla_\xi^2)^M P(\mathbf{x}, \boldsymbol{\xi})\phi(\mathbf{x})(1-\psi(\mathbf{y}))|\mathbf{x}-\mathbf{y}|^{-2M} e^{i(\mathbf{x}-\mathbf{y})\cdot\boldsymbol{\xi}}u(\mathbf{y})$$

using the identity

$$e^{i(\mathbf{x}-\mathbf{y})\cdot\boldsymbol{\xi}} = |\mathbf{x}-\mathbf{y}|^{-2}\left[-\nabla_\xi^2 e^{i(\mathbf{x}-\mathbf{y})\cdot\boldsymbol{\xi}}\right]$$

and integrating by parts $2M$ times in $\boldsymbol{\xi}$. This is permissible because $\phi(\mathbf{x})(1-\psi(\mathbf{y})) \neq 0 \Rightarrow |\mathbf{x}-\mathbf{y}| > \delta$.

According to the definition of ΨDO,

$$|(-\nabla_\xi^2)^M P(\mathbf{x}, \boldsymbol{\xi})| \leq C|\boldsymbol{\xi}|^{m-2M}$$

For any K, the integral thus becomes absolutely convergent after K differentiations of the integrand, provided M is chosen large enough. Q.E.D. Claim.

This leaves us with $\phi P(\mathbf{x}, D)(\psi u)$. Pick $\eta \in \Xi'$ and w.l.o.g. scale $|\eta| = 1$.

Fourier transform:

$$\mathcal{F}(\phi P(\mathbf{x}, D)(\psi u))(\tau\eta) = \int dx \int d\xi\, P(\mathbf{x}, \boldsymbol{\xi})\phi(\mathbf{x})\hat{\psi u}(\xi)e^{i\mathbf{x}\cdot(\boldsymbol{\xi}-\tau\eta)}$$

Introduce $\tau\theta = \xi$, and rewrite this as

$$= \tau^n \int dx \int d\theta\, P(\mathbf{x}, \tau\theta)\phi(\mathbf{x})\hat{\psi u}(\tau\theta)e^{i\tau\mathbf{x}\cdot(\theta-\eta)}$$

Divide the domain of the inner integral into $\{\theta : |\theta - \eta| > \epsilon\}$ and its complement. Use

$$-\nabla_x^2 e^{i\tau\mathbf{x}\cdot(\theta-\eta)} = \tau^2|\theta - \eta|^2 e^{i\tau\mathbf{x}\cdot(\theta-\eta)}$$

Integrate by parts $2M$ times to estimate the first integral:

$$\tau^{n-2M} \left| \int dx \int_{|\theta-\eta|>\epsilon} d\theta \, (-\nabla_x^2)^M [P(\mathbf{x}, \tau\theta)\phi(\mathbf{x})] \hat{\psi u}(\tau\theta) \right.$$

$$\left. \times \, |\theta - \eta|^{-2M} e^{i\tau\mathbf{x}\cdot(\theta-\eta)} \right|$$

$$\leq C\tau^{n+m-2M}$$

m being the order of P. Thus the first integral is rapidly decreasing in τ.

For the second integral, note that $|\theta - \eta| \leq \epsilon \Rightarrow \theta \in \Xi$, per the defn of Ξ'. Since $X \times \Xi$ is disjoint from the wavefront set of u, for a sequence of constants C_N, $|\hat{\psi u}(\tau\theta)| \leq C_N \tau^{-N}$ uniformly for θ in the (compact) domain of integration, whence the second integral is also rapidly decreasing in τ. **Q. E. D.**

And that's why Kirchhoff migration works, at least in the simple geometric optics regime.

Inversion aperture

$\Gamma[v] \subset \mathbf{R}^3 \times \mathbf{R}^3 - 0$:

if $WF(r) \subset \Gamma[v]$, then $WF(F[v]^* F[v] r) = WF(r)$ and $F[v]^* F[v]$ "acts invertible". [construction of $\Gamma[v]$ - later!]

Beylkin: with proper choice of amplitude $b(\mathbf{x}_r, t; \mathbf{x}_s)$, the modified Kirchhoff migration operator

$$F[v]^\dagger d(\mathbf{x}) = \int\int\int dx_r\, dx_s\, dt\, b(\mathbf{x}_r, t; \mathbf{x}_s)\delta(t - \tau(\mathbf{x}; \mathbf{x}_s) - \tau(\mathbf{x}; \mathbf{x}_r))d(\mathbf{x}_r, t; \mathbf{x}_s)$$

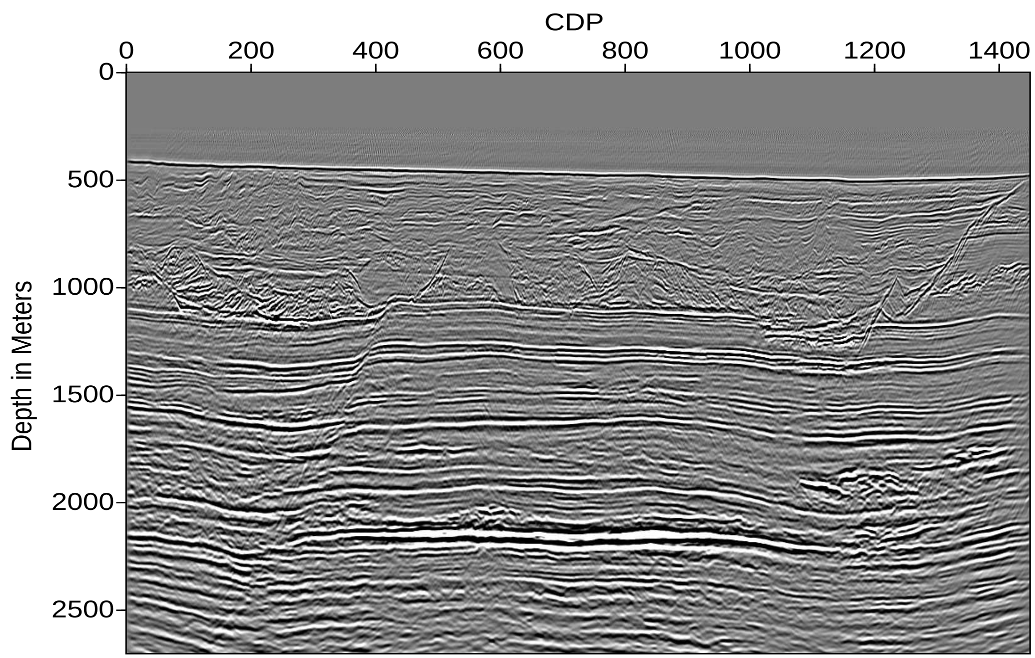
yields $F[v]^\dagger F[v] r \simeq r$ if $WF(r) \subset \Gamma[v]$

For details of Beylkin construction: Beylkin, 1985; Miller et al 1989; Bleistein, Cohen, and Stockwell 2000; WWS Math Foundations, MGSS notes 1998. All components are by-products of eikonal solution.

aka: Generalized Radon Transform ("GRT") inversion, Ray-Born inversion, migration/inversion, true amplitude migration,...

Many extensions, eg. to elasticity: Bleistein, Burridge, deHoop, Lambaré,...

Apparent limitation: construction relies on simple geometric optics (no multipathing) - is this really necessary?



Example of GRT Inversion (application of $F[v]^\dagger$): K. Araya (1995), "2.5D" inversion of marine streamer data from Gulf of Mexico: 500 source positions, 120 receiver channels, 750 Mb.

Why Beylkin isn't enough

The theory developed by Beylkin and others cannot be the end of the story:

- The "single ray" hypotheses generally fails in the presence of strong refraction.
- B. White, "The Stochastic Caustic" (1982): For "random but smooth" $v(\mathbf{x})$ with variance σ, points at distance $O(\sigma^{-2/3})$ from source have more than one ray connecting to source, with probability 1 - *multipathing* associated with formation of *caustics* = ray envelopes.
- Formation of caustics invalidates asymptotic analysis on which Beylkin result is based.

Why it matters

- Strong refraction leading to multipathing and caustic formation typical of salt (4-5 km/s) intrusion into sedimentary rock (2-3 km/s) (eg. Gulf of Mexico), also chalk tectonics in North Sea and elsewhere - some of the most promising petroleum provinces!

Escape from simplicity - the Canonical Relation

How do we get away from "simple geometric optics", SSR, DSR,... - all violated in sufficiently complex (and realistic) models? Rakesh *Comm. PDE* 1988, Nolan *Comm. PDE* 1997: global description of $F_\delta[v]$ as mapping reflectors $\mapsto$ reflections.

$Y = \{\mathbf{x}_s, t, \mathbf{x}_r\}$ (time $\times$ set of source-receiver pairs) submfd of $\mathbf{R}^7$ of dim. ≤ 5, $\Pi : T^*(\mathbf{R}^7) \to T^*Y$ the natural projection

$\mathrm{supp}\, r \subset X \subset \mathbf{R}^3$

Canonical relation $C_{F_\delta[v]} \subset T^*(X) - \{\mathbf{0}\} \times T^*(Y) - \{\mathbf{0}\}$ describes singularity mapping properties of F:

$$(\mathbf{x}, \xi, \mathbf{y}, \eta) \in C_{F_\delta[v]} \Leftrightarrow$$

$$\text{for some } u \in \mathcal{E}'(X),\ (\mathbf{x}, \xi) \in WF(u),\ \text{and } (\mathbf{y}, \eta) \in WF(Fu)$$

Rays Construction of the Relation

Rays of geometric optics: solutions of Hamiltonian system

$$\frac{d\mathbf{X}}{dt} = \nabla_{\mathbf{\Xi}} H(\mathbf{X}, \mathbf{\Xi}), \;\; \frac{d\mathbf{\Xi}}{dt} = -\nabla_{\mathbf{X}} H(\mathbf{X}, \mathbf{\Xi})$$

with $H(\mathbf{X}, \mathbf{\Xi}) = \frac{1}{2}[1 - v^2(\mathbf{X})|\mathbf{\Xi}|^2] = 0$ (*null bicharacteristics*).

Characterization of C_F:

$$((\mathbf{x}, \xi), (\mathbf{x}_s, t, \mathbf{x}_r, \xi_\mathbf{s}, \tau, \xi_\mathbf{r})) \in C_{F_\delta[v]} \subset T^*(X) - \{\mathbf{0}\} \times T^*(Y) - \{\mathbf{0}\}$$

$\Leftrightarrow$ there are *rays of geometric optics* $(\mathbf{X}_s, \mathbf{\Xi}_s)$, $(\mathbf{X}_r, \mathbf{\Xi}_r)$ and times t_s, t_r so that

$$\Pi(\mathbf{X}_s(0), t, \mathbf{X}_r(t), \mathbf{\Xi}_s(0), \tau, \mathbf{\Xi}_r(t)) = (\mathbf{x}_s, t, \mathbf{x}_r, \xi_s, \tau, \xi_r),$$

$$\mathbf{X}_s(t_s) = \mathbf{X}_r(t - t_r) = \mathbf{x}, \; t_s + t_r = t, \; \mathbf{\Xi}_s(t_s) - \mathbf{\Xi}_r(t - t_r) || \xi$$

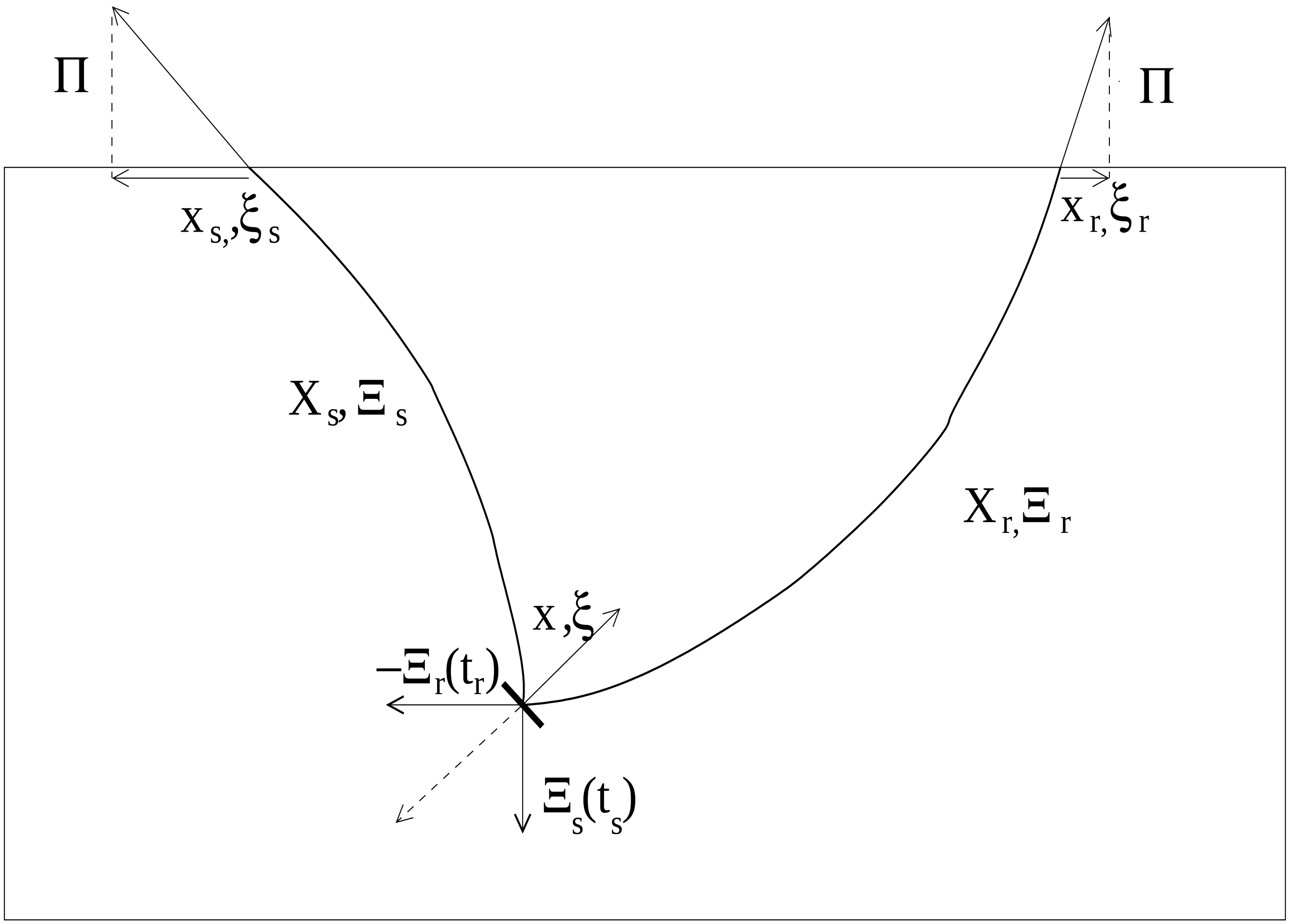
Since $\boldsymbol{\Xi}_s(t_s)$, $-\boldsymbol{\Xi}_r(t - t_r)$ have same length, sum = bisector $\Rightarrow$ *velocity vectors of incident ray from source and reflected ray from receiver (traced backwards in time) make equal angles with reflector at* $\mathbf{x}$ *with normal* ξ.

Upshot: canonical relation of $F_\delta[v]$ simply enforces the equal-angles law of reflection.

Further, *rays carry high-frequency energy*, in exactly the fashion that seismologists imagine.

Finally, *Rakesh's characterization of* C_F *is global:* no assumptions about ray geometry, other than no forward scattering and no grazing incidence on the acquisition surface Y, are needed.

The Picture



Proof: Plan of attack

Recall that

$$F[v]r(\mathbf{x}_r, t; \mathbf{x}_s) = \frac{\partial \delta u}{\partial t}(\mathbf{x}_r, t; \mathbf{x}_s)$$

where

$$\frac{1}{v^2}\frac{\partial^2 \delta u}{\partial t^2} - \nabla^2 \delta u = \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} r$$

$$\frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} - \nabla^2 u = \delta(t)\delta(\mathbf{x} - \mathbf{x}_s)$$

and $u, \delta u \equiv 0,\ t < 0$.

Need to understand (1) $WF(u)$, (2) relation $WF(r) \leftrightarrow WF(ru)$, (3) WF of soln of WE in terms of WF of RHS (this also gives (1)!).

Singularities of the Acoustic Potential Field

Main tool: **Propagation of Singularities** theorem of Hörmander (1970).

Given symbol $p(\mathbf{x},\boldsymbol{\xi})$, order m, with asymptotic expansion, define *null bicharateristics* (= rays) as solutions $(\mathbf{x}(t),\boldsymbol{\xi}(t))$ of Hamiltonian system

$$\frac{d\mathbf{x}}{dt}=\frac{\partial p}{\partial \boldsymbol{\xi}}(\mathbf{x},\boldsymbol{\xi}),\ \ \frac{d\boldsymbol{\xi}}{dt}=-\frac{\partial p}{\partial \mathbf{x}}(\mathbf{x},\boldsymbol{\xi})$$

with $p(\mathbf{x}(t),\boldsymbol{\xi}(t))\equiv 0$.

Theorem: Suppose $p(\mathbf{x},D)u=f$, and suppose that for $t_0\le t\le t_1$, $(\mathbf{x}(t),\boldsymbol{\xi}(t))\notin WF(f)$. Then either $\{(\mathbf{x}(t),\boldsymbol{\xi}(t)) : t_0\le t\le t_1\}\subset WF(u)$ or $\{(\mathbf{x}(t),\boldsymbol{\xi}(t)) : t_0\le t\le t_1\}\subset T^*(\mathbf{R}^n)-WF(u)$.

Source to Field

RHS of wave equation for u = δ function in $\mathbf{x}, t$. WF set = $\{(\mathbf{x}, t, \boldsymbol{\xi}, \tau) : \mathbf{x} = \mathbf{x}_s, t = 0\}$ - i.e. no restriction on covector part.

$\Rightarrow$ $(\mathbf{x}, t, \boldsymbol{\xi}, \tau) \in WF(u)$ iff a ray starting at $(\mathbf{x}_s, 0)$ passes over $(\mathbf{x}, t)$ - i.e. $(\mathbf{x}, t)$ lies on the "light cone" with vertex at $(\mathbf{x}_x, 0)$. Symbol for wave op is $p(\mathbf{x}, t, \boldsymbol{\xi}, \tau) = \frac{1}{2}(\tau^2 - v^2(\mathbf{x})|\boldsymbol{\xi}|^2)$, so Hamilton's equations for null bicharacteristics are

$$\frac{d\mathbf{X}}{dt} = -v^2(\mathbf{X})\boldsymbol{\Xi}, \ \ \frac{d\boldsymbol{\Xi}}{dt} = \nabla \log v(\mathbf{X})$$

Thus $\boldsymbol{\xi}$ is proportional to velocity vector of ray.

[$(\boldsymbol{\xi}, \tau)$ *normal* to light cone.]

Singularities of Products

To compute $WF(ru)$ from $WF(r)$ and $WF(u)$, use *Gabor calculus* (Duistermaat, Ch. 1)

Here r is really $(r \circ \pi)u$, where $\pi(\mathbf{x}, t) = \mathbf{x}$. Choose bump function ϕ localized near $(\mathbf{x}, t)$

$$\phi\widehat{(r \circ \pi)}u(\boldsymbol{\xi}, \tau) = \int d\boldsymbol{\xi}' \, d\tau' \widehat{\phi r}(\boldsymbol{\xi}') \delta(\tau') \widehat{u}(\boldsymbol{\xi} - \boldsymbol{\xi}', \tau - \tau')$$

$$= \int d\boldsymbol{\xi}' \widehat{\phi r}(\boldsymbol{\xi}') \widehat{u}(\boldsymbol{\xi} - \boldsymbol{\xi}', \tau)$$

This will decay rapidly as $|(\boldsymbol{\xi}, \tau)| \to \infty$ unless (i) you can find $(\mathbf{x}', \boldsymbol{\xi}') \in WF(r)$ so that $\mathbf{x}, \mathbf{x}' \in \pi(\mathrm{supp}\phi)$, $\boldsymbol{\xi} - \boldsymbol{\xi}' \in WF(u)$, i.e. $(\boldsymbol{\xi}, \tau) \in WF(r \circ \pi) + WF(u)$, or (ii) $\boldsymbol{\xi} \in WF(r)$ or $(\boldsymbol{\xi}, \tau) \in WF(u)$.

Possibility (ii) will not contribute, so effectively

$$WF((r \circ \pi)u) = \{(\mathbf{x}, t_s, \boldsymbol{\xi} + \boldsymbol{\Xi}_s(t_s), \cdot) : (\mathbf{x}, \boldsymbol{\xi}) \in WF(r),\ \mathbf{x} = \mathbf{X}_s(t_s)$$

for a ray $(\mathbf{X}_s, \boldsymbol{\Xi}_s)$ with $\mathbf{X}_s(0) = x_s$, some τ.

Wavefront set of Scattered Field

Once again use propagation of singularities: $(\mathbf{x}_r, t, \boldsymbol{\xi}_r, \tau_r) \in WF(\delta u) \Leftrightarrow$ on ray $(\mathbf{X}_r, \boldsymbol{\Xi}_r)$ passing through $WF(ru)$. Can argue that time of intersection is $t - t_r < t$.

That is,

$$\mathbf{X}_r(t) = \mathbf{x}_r, \mathbf{X}_r(t - t_r) = \mathbf{X}_s(t_s) = x,$$

$t = t_r + t_s$, and

$$\boldsymbol{\Xi}_r(t_s) = \boldsymbol{\xi} + \boldsymbol{\Xi}_s(t_s)$$

for some $\boldsymbol{\xi} \in WF(r)$. **Q. E. D.**

Rakesh's Thesis

Rakesh also showed that $F[v]$ is a *Fourier Integral Operator* = class of oscillatory integral operators, introduced by Hörmander and others in the '70s to describe the solutions of nonelliptic PDEs.

Phases and amplitudes of FIOs satisfy certain restrictive conditions. Canonical relations have geometric description similar to that of $F[v]$. Adjoint of FIO is FIO with inverse canonical relation.

ΨDOs are special FIOs, as are GRTs.

Composition of FIOs does *not* yield an FIO in general. Beylkin had shown that $F[v]^*F[v]$ is FIO (ΨDO, actually) under simple ray geometry hypothesis - but this is only sufficient. Rakesh noted that this follows from general results of Hörmander: *simple ray geometry* $\Leftrightarrow$ *canonical relation is graph of ext. deriv. of phase function*.

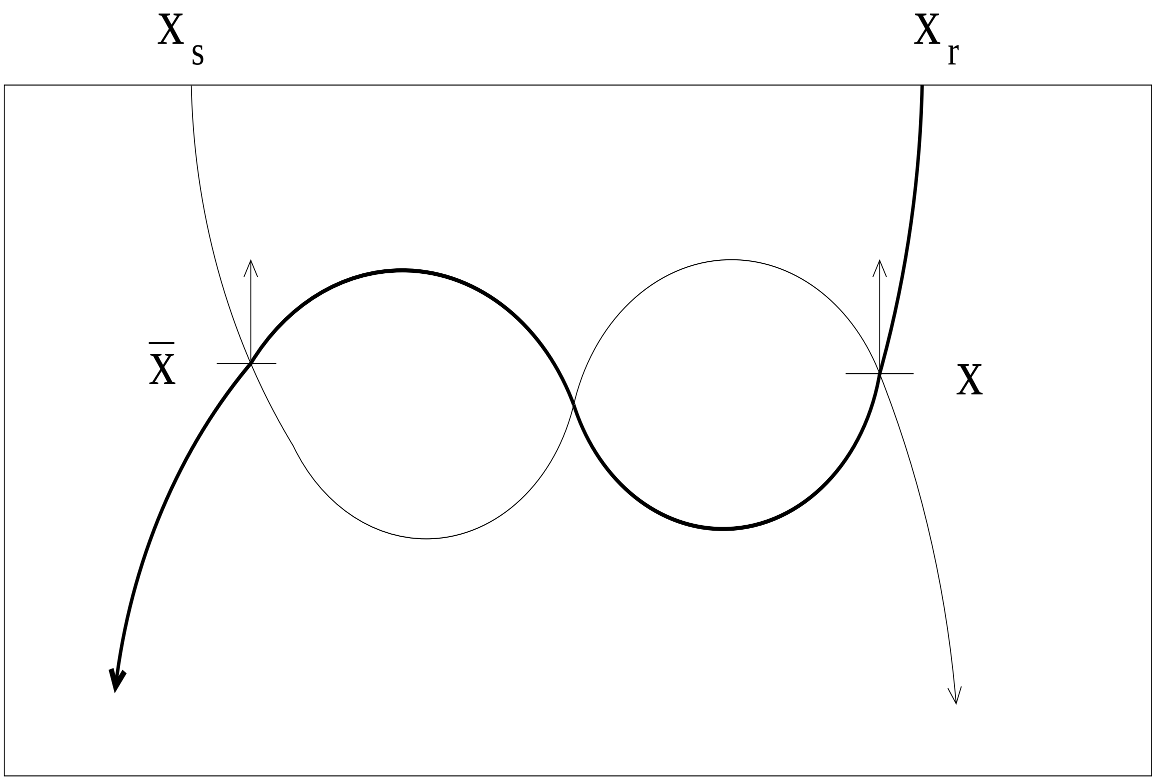
The Shell Guys and TIC

Smit, tenKroode and Verdel (1998): provided that

- source, receiver positions $(\mathbf{x}_s, \mathbf{x}_r)$ form an *open* 4D manifold ("complete coverage" - all source, receiver positions at least locally), and
- the *Traveltime Injectivity Condition* ("TIC") holds: $C_{F[v]}^{-1} \subset T^*Y - \{0\} \times T^*X - \{0\}$ is a *function* - that is, initial data for source and receiver rays and total travel time together determine reflector uniquely.

then $F[v]^* F[v]$ is ΨDO $\Rightarrow$ application of $F[v]^*$ produces image, and $F[v]^* F[v]$ has microlocal parametrix ("asymptotic inversion").

TIC is a nontrivial constraint!



Symmetric waveguide: time ($\mathbf{x}_s \rightarrow \bar{\mathbf{x}} \rightarrow \mathbf{x}_r$) same as time ($\mathbf{x}_s \rightarrow \mathbf{x} \rightarrow \mathbf{x}_r$), so TIC fails.

Stolk's Thesis

Stolk (2000): under "complete coverage" hypothesis, v for which $F[v]^*F[v]$ is = [ΨDO + rel. smoothing op] form open, dense set (without assuming TIC!).

NB: application of $F[v]^*$ involves accounting for *all* rays connecting source and receiver with reflectors. Standard practice still attempts imaging with single choice of ray pair (shortest time, max energy,...). Operto et al (2000) give nice illustration that all rays must be included.

Nolan's Thesis

Limitation of Smit-tenKroode-Verdel: most idealized data acquisition geometries violate "complete coverage": for example, idealized marine streamer geometry (src-recvr submfd is 3D)

Nolan (1997): result remains true without "complete coverage" condition: requires only TIC plus addl condition so that projection $C_{F[v]} \to T^*Y$ is embedding - but examples violating TIC are much easier to construct when source-receiver submfd has positive codim.

Sinister Implication: When data is just a single gather - common shot, common offset - image may contain *artifacts*, i.e. spurious reflectors not present in model.

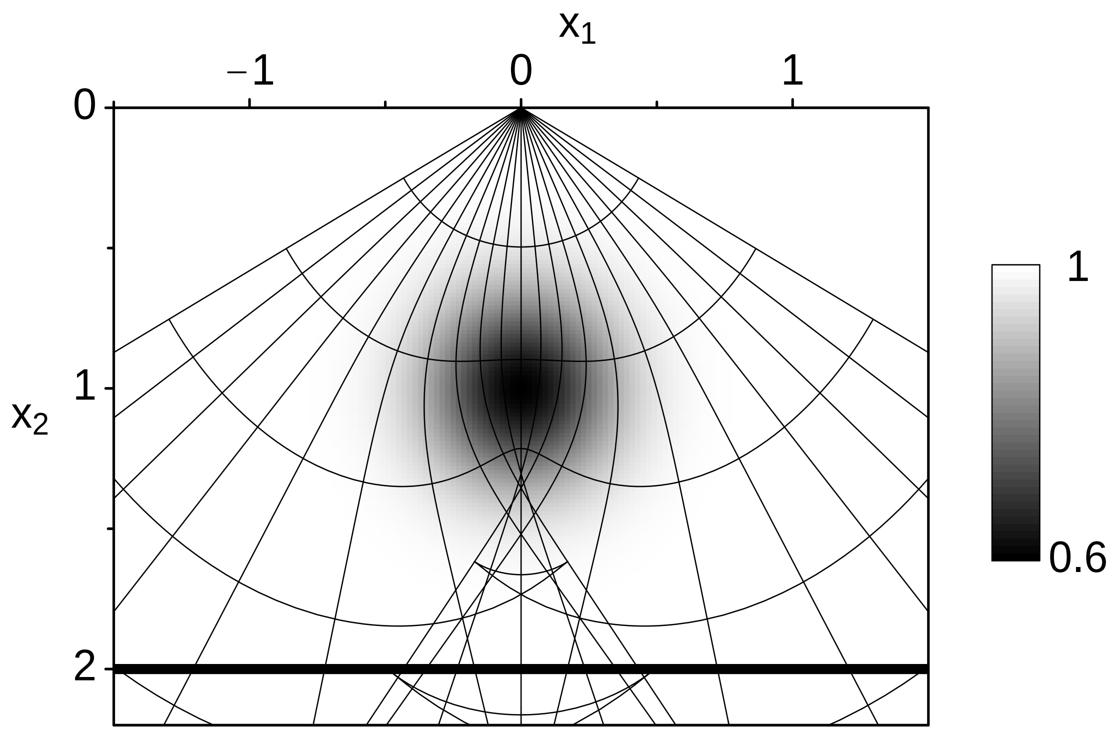
Horrible Example I

Synthetic 2D Example (Stolk and WWS, 2001 - *Geophysics* 2004)

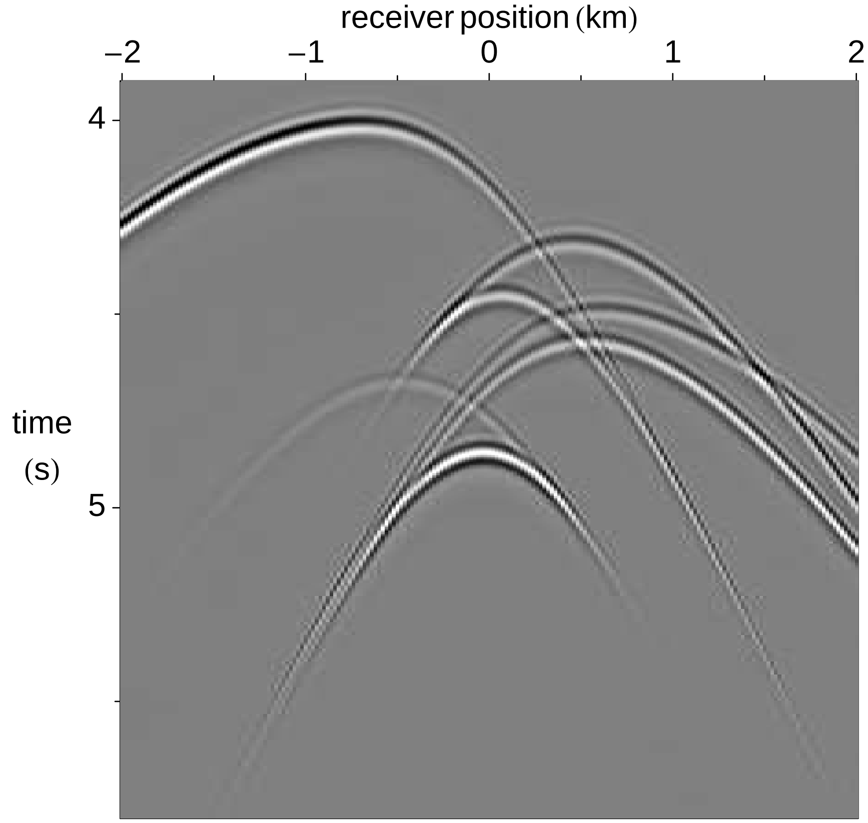
Strongly refracting acoustic lens (v) over horizontal reflector (r), $d = F[v]r$.

(i) for open source-receiver set, $F[v]^*d$ = good image of reflector - within limits of finite frequency implied by numerical method, $F[v]^*F[v]$ acts like ΨDO;

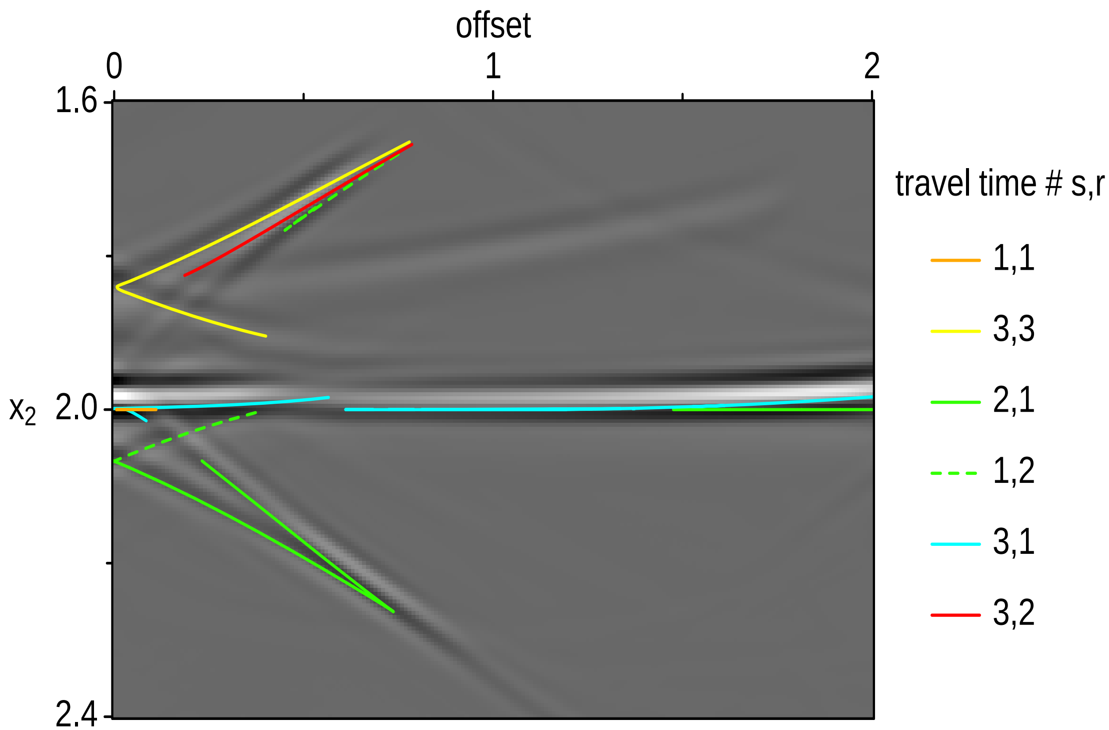
(ii) for *common offset* submfd (codim 1), TIC is violated and $WF(F[v]^*d)$ is larger than $WF(r)$.



Gaussian lens velocity model, flat reflector at depth 2 km, overlain with rays and wavefronts (Stolk & S. 2002 SEG).



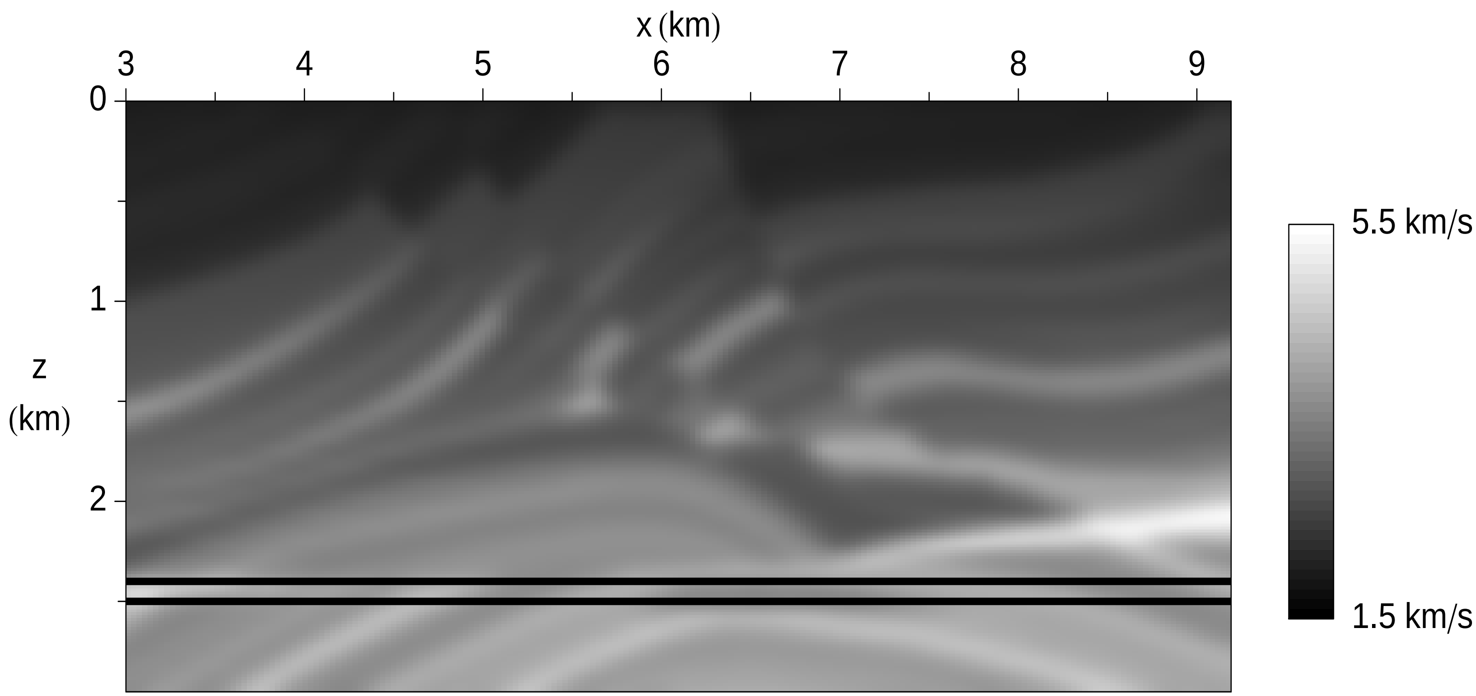
Typical shot gather - lots of arrivals

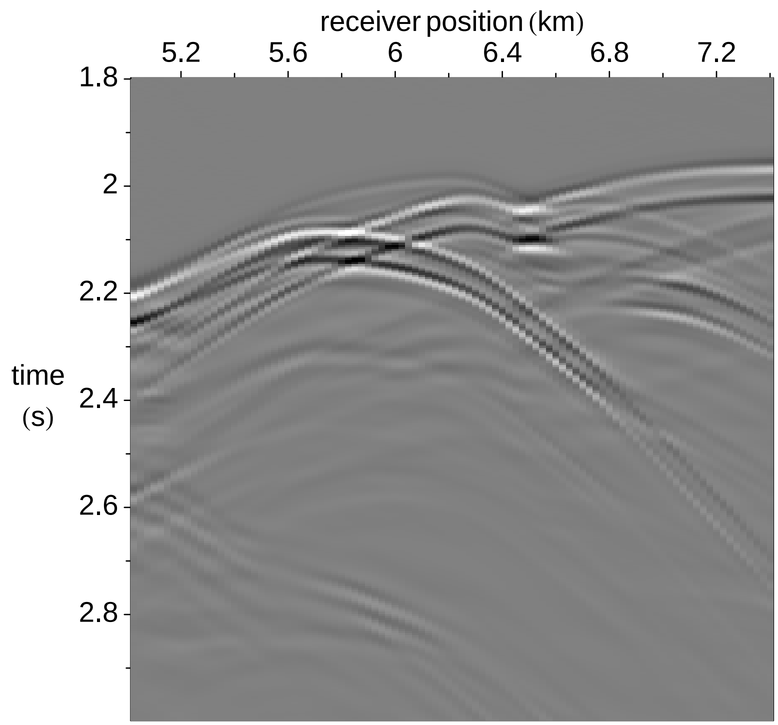


Offset common image gather (slice of $\tilde{F}[v]^*d$), with kinematically predicted reflector images overlain.

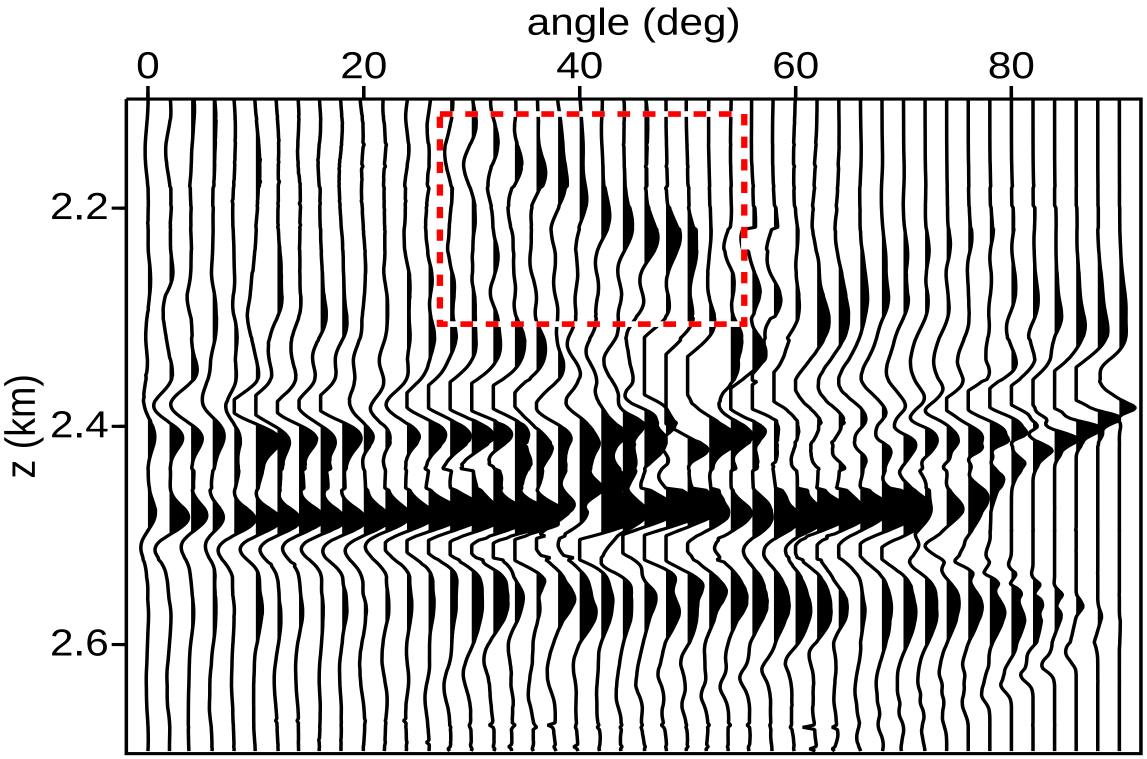
Horrible Example II

Stolk and Symes, *Geophysics* 2004: “Marmouflat” model = smoothed Marmousi (Versteeg & Grau 1991) with two flat reflectors.

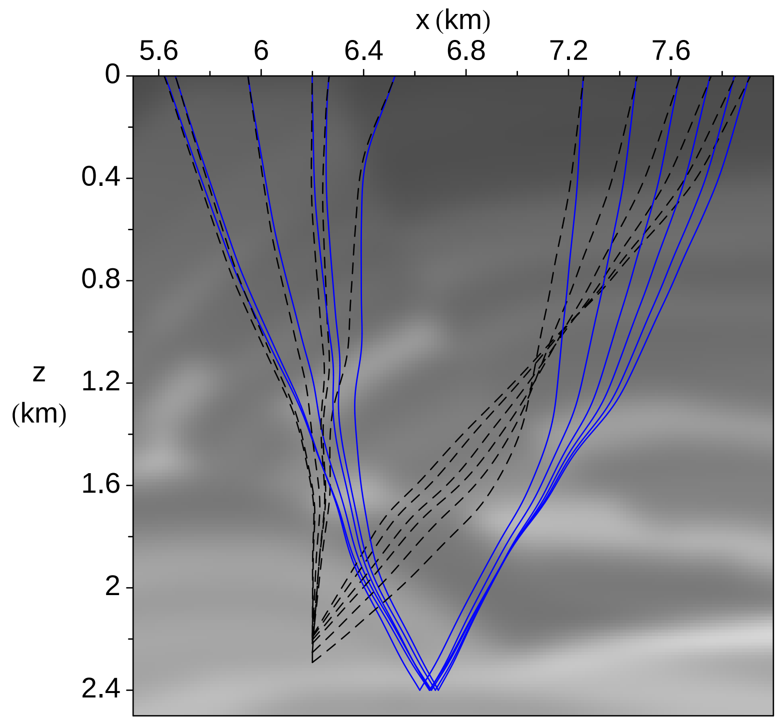




Typical shot gather: much evidence of multipathing, caustic formation.



Typical common scattering angle image gather: note nonflat event in box - results from data event migrating along *wrong ray pair*.



Blue rays = energy path producing data event. Black rays: energy path for migration, resulting in displaced, angle-dependent image artifact.

What it all means

Note that a gather scheme makes the scattering operator block-diagonal: for example with data sorted into common offset gathers $h = (x_r - x_s)/2$,

$$F[v] = [F_{h_1}[v], ..., F_{h_N}[v]]^T, \; d = [d_{h_1}, ..., d_{h_N}]^T$$

Thus $F[v]^* d = \sum_i F_{h_i}[v]^* d_{h_i}$. Otherwise put: to form image, **migrate** ith gather (apply migration operator $F_{h_i}[v]^*$, then **stack** individual migrated images (hence *prestack migration*).

Horrible Examples show that individual migrated images may contain nonphysical apparent reflectors (artifacts).

Smit-tenKroode-Verdel, Nolan, Stolk: if TIC holds, then these artifacts are not stationary with respect to the gather parameter, hence *stack out* (interfere destructively) in final image.

Mathematics of Seismic Imaging Part II - addendum on Wave Equation Migration

William W. Symes

PIMS, July 2005

Wave Equation Migration

Techniques for computing $F[v]^*$:

(i) Reverse time

(ii) Reverse depth

Reverse Time Migration, Zero Offset

Start with the zero-offset case - easier, but only if you replace it with the exploding reflector model, which replaces $F[v]$ by

$$\tilde{F}[v]r(\mathbf{x}_s, t) = w(\mathbf{x}_s, t),\ \ \mathbf{x}_s \in X_s, 0 \le t \le T$$

$$\left(\frac{4}{v^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) w = \delta(t)\frac{2r}{v^2},\ \ w \equiv 0,\ t < 0$$

To compute the adjoint, start with its definition: choose $d \in \mathcal{E}(X_s \times (0, T))$, so that

$$< \tilde{F}[v]^* d, r >=< d, \tilde{F}[v]r >$$

$$= \int_{X_s} dx_s \int_0^T dt\, d(\mathbf{x}_s, t) w(\mathbf{x}_s, t)$$

The only thing you know about w is that it solves a wave equation with r on the RHS. To get this fact into play, (i) rewrite the integral as a space-time integral:

$$= \int_{\mathbf{R}^3} dx \int_0^T dt \int_{X_s} dx_s \, d(\mathbf{x}_s, t)\delta(\mathbf{x} - \mathbf{x}_s)w(\mathbf{x}, t)$$

(ii) write the other factor in the integrand as the image of a field q under the (adjoint of the) wave operator (it's self-adjoint), that is,

$$\left(\frac{4}{v^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) q(\mathbf{x}, t) = \int_{X_s} dx_s \, d(\mathbf{x}_s, t)\delta(\mathbf{x} - \mathbf{x}_s)$$

so

$$= \int_{\mathbf{R}^3} dx \int_0^T dt \left[\left(\frac{4}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right) q(\mathbf{x}, t)\right] w(\mathbf{x}, t)$$

(iii) integrate by parts

$$= \int_{\mathbf{R}^3} dx \int_0^T dt \left[\left(\frac{4}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right) w(\mathbf{x},t)\right] q(\mathbf{x},t)$$

which works if $q \equiv 0,\ t > T$ (*final value condition*); (iv) use the wave equation for w

$$= \int_{\mathbf{R}^3} dx \int_0^T dt \, \frac{2}{v(\mathbf{x})^2} r(\mathbf{x})\delta(t) q(\mathbf{x},t)$$

(v) observe that you have computed the adjoint:

$$= \int_{\mathbf{R}^3} dx \, r(\mathbf{x}) \left[\frac{2}{v(\mathbf{x})^2} q(\mathbf{x},0)\right] =< r, \tilde{F}[v]^* d >$$

i.e.

$$\tilde{F}[v]^* d = \frac{2}{v(\mathbf{x})^2} q(\mathbf{x},0)$$

Summary of the computation, with the usual description:

- Use that data as sources, backpropagate in time - i.e. solve the final value ("reverse time") problem

$$\left(\frac{4}{v^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) q(\mathbf{x},t) = \int_{X_s} dx_s\, d(\mathbf{x}_s,t)\delta(\mathbf{x}-\mathbf{x}_s),\;\; q \equiv 0,\; t > T$$

- read out the "image" (= adjoint output) at $t = 0$:

$$\tilde{F}[v]^* d = \frac{2}{v(\mathbf{x})^2} q(\mathbf{x},0)$$

Note: The adjoint (time-reversed) field q is *not* the physical field (δu) run backwards in time, contrary to some imputations in the literature.

Historical Remarks

- Known as "two way reverse time finite difference poststack migration" in geophysical literature (Whitmore, 1982)
- uses full (two way) wave equation, propagates adjoint field backwards in time, generally implemented using finite difference discretization.
- Same as "adjoint state method", Lions 1968, Chavent 1974 for control and inverse problems for PDEs - much earlier for control of ODEs - Lailly, Tarantola '80s.
- My buddy Tapia says: all you're doing is transposing a matrix! True (after discretization), but it's important that these matrices are triangular, so can be implemented by recursions - forward for simulation, backwards for adjoint.

Reverse Time Migration, Prestack

A slightly messier computation computes the adjoint of $F[v]$ (i.e. multioffset or *prestack* migration):

$$F[v]^* d(\mathbf{x}) = -\frac{2}{v(\mathbf{x})} \int dx_s \int_0^T dt \left(\frac{\partial q}{\partial t} \nabla^2 u \right) (\mathbf{x}, t; \mathbf{x}_s)$$

where *adjoint field* q satisfies $q \equiv 0$, $t \geq T$ and

$$\left(\frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) q(\mathbf{x}, t; \mathbf{x}_s) = \int dx_r \, d(\mathbf{x}_r, t; \mathbf{x}_s) \delta(\mathbf{x} - \mathbf{x}_r)$$

Proof

$$< F[v]^*d, r > = < d, F[v]r >$$

$$= \int \int dx_s \, dx_r \int_0^T dt \, d(\mathbf{x}_r, t; \mathbf{x}_s) \frac{\partial \delta u}{\partial t}(\mathbf{x}_r, t; \mathbf{x}_s)$$

$$= \int dx_s \int dx \int_0^T dt \left\{ \int dx_r \, d(\mathbf{x}_r, t; \mathbf{x}_s) \delta(\mathbf{x} - \mathbf{x}_r) \right\} \frac{\partial \delta u}{\partial t}(\mathbf{x}, t; \mathbf{x}_s)$$

$$= \int dx_s \int dx \int_0^T dt \left[\left(\frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) q \right] \frac{\partial \delta u}{\partial t}(\mathbf{x}, t; \mathbf{x}_s)$$

$$= -\int dx_s \int dx \int_0^T dt \left[\left(\frac{1}{v^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\delta u\right]\frac{\partial q}{\partial t}(\mathbf{x}, t; \mathbf{x}_s)$$

(boundary terms in integration by parts vanish because (i) $\delta u \equiv 0,\ t << 0$; (ii) $q \equiv 0,\ t >> 0$; (iii) both vanish for large $\mathbf{x}$, at each t)

$$= -\int dx_s \int dx \int_0^T dt \left(\frac{2r}{v^2}\frac{\partial^2 u}{\partial t^2}\frac{\partial q}{\partial t}\right)(\mathbf{x}, t; \mathbf{x}_s)$$

$$= -\int dx_s \int dx\, r(\mathbf{x})\frac{2}{v^2(\mathbf{x})}\int_0^T dt \left(\frac{\partial^2 u}{\partial t^2}\frac{\partial q}{\partial t}\right)(\mathbf{x}, t; \mathbf{x}_s)$$

$$=< r, F[v]^* d >$$

q.e.d.

Implementation

Algorithm: finite difference or finite element discretization in $\mathbf{x}$, finite difference time stepping.

- For each $\mathbf{x}_s$, solve wave equation for u forward in t, record final (t=T) Cauchy data, also (for example) Dirichlet boundary data.
- Step u and q backwards in time together; at each time step, data serves as source for q ("backpropagate data")
- During backwards time stepping, accumulate (approximations to)

$$Q(\mathbf{x}) += \frac{2}{v^2(\mathbf{x})} \int_0^T dt \left(\frac{\partial^2 u}{\partial t^2} \frac{\partial q}{\partial t} \right) (\mathbf{x}, t; \mathbf{x}_s)$$

  ("crosscorrelate reference and backpropagated field").
- next $\mathbf{x}_s$ - after last $\mathbf{x}_s$, $F[v]^* d = Q$.

Reverse Depth Migration, Zero Offset

aka: depth extrapolation, downward continuation, or simply "wave equation migration".

Introduced by Claerbout, early 70's ("swimming pool equation"). Again, assume exploding reflector model:

$$\tilde{F}[v]r(\mathbf{x}_s, t) = w(\mathbf{x}_s, t),\ \mathbf{x}_s \in X_s, 0 \le t \le T$$

$$\left(\frac{4}{v^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) w = \delta(t)\frac{2r}{v^2},\ w \equiv 0,\ t < 0$$

Basic idea: 2nd order wave equation permits waves to move in all directions, but waves carrying reflected energy are (mostly) moving *up*. Should satisfy a 1st order equation for wave motion in one direction.

Coming up...

For the moment use 2D notation $\mathbf{x} = (x, z)$ etc. Write wave equation as evolution equation in z:

$$\frac{\partial^2 w}{\partial z^2} - \left(\frac{4}{v^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}\right) w = -\delta(t)\frac{2r}{v^2}$$

Suppose that you could take the square root of the operator in parentheses - call it B. Then the LHS of the wave equation becomes

$$\left(\frac{\partial}{\partial z} - B\right)\left(\frac{\partial}{\partial z} + B\right) w = -\delta(t)\frac{2r}{v^2}$$

so setting $\tilde{w} = \left(\frac{\partial}{\partial z} + B\right) w$ you get

$$\left(\frac{\partial}{\partial z} - B\right) \tilde{w} = -\delta(t)\frac{2r}{v^2}$$

Some issues

This *might* be the required equation for upcoming waves.

Two major problems: (i) how the h–l do you take the square root of a PDO?

(ii) what guarantees that the equation just written governs upcoming waves?

Answers to be found in the theory of ΨDOs!

Classical ΨDOs

Important *subclass* of *classical* ΨDOs: those whose ("classical") symbols have asymptotic expansions:

$$p(\mathbf{x}, \boldsymbol{\xi}) \sim \sum_{j \leq m} p_j(\mathbf{x}, \boldsymbol{\xi}), \ |\boldsymbol{\xi}| \rightarrow \infty$$

in which p_j is *homogeneous in* $\boldsymbol{\xi}$ *of degree* j:

$$p_j(\mathbf{x}, \tau\boldsymbol{\xi}) = \tau^j p_j(\mathbf{x}, \tau\boldsymbol{\xi}), \ \tau, |\boldsymbol{\xi}| \geq 1$$

The *principal symbol* is the homogeneous term of highest degree, i.e. p_m above.

Products of ΨDOs are ΨDOs.

Classical ΨDOs have more complete *calculus*, including prescriptions for "computing" adjoints, products, and the like. From now on unless otherwise stated, all ΨDOs are classical.

Product rule for ΨDOs: if p^1, p^2 are classical,

$$p^1(\mathbf{x},\boldsymbol{\xi}) = \sum_{j\le m^1} p^1_j(\mathbf{x},\boldsymbol{\xi}),\ \ p^2(\mathbf{x},\boldsymbol{\xi}) = \sum_{j\le m^2} p^2_j(\mathbf{x},\boldsymbol{\xi})$$

then so is $p^1(\mathbf{x},D)p^2(\mathbf{x},D)$, and its principal symbol is $p^1_{m^1}(\mathbf{x},\boldsymbol{\xi})p^2_{m^2}(\mathbf{x},\boldsymbol{\xi})$, and there is an algorithm for computing the rest of the expansion.

In an open neighborhood $X \times \Xi$ of $(\mathbf{x}_0,\boldsymbol{\xi}_0)$, symbol of $p^1(\mathbf{x},D)p^2(\mathbf{x},D)$ depends only on symbols of p^1, p^2 in $X \times \Xi$.

Consequence: if $a(\mathbf{x}, D)$ has an asymptotic expansion and is of order $m \in \mathbf{R}$, and $a_m(\mathbf{x}_0, \boldsymbol{\xi}_0) > 0$ in $\mathcal{P} \subset \mathbf{R}^n \times \mathbf{R}^n - 0$, then there exists $b(\mathbf{x}, D)$ of order $m/2$ with asymptotic expansion for which

$$(a(\mathbf{x}, D) - b(\mathbf{x}, D)b(\mathbf{x}, D))u \in \mathcal{E}(\mathbf{R}^n)$$

for any $u \in \mathcal{E}'(\mathbf{R}^n)$ with $WF(u) \subset \mathcal{P}$.

Moreover, $b_{m/2}(\mathbf{x}, \boldsymbol{\xi}) = \sqrt{a_m(\mathbf{x}, \boldsymbol{\xi})},\ (\mathbf{x}, \boldsymbol{\xi}) \in \mathcal{P}$. Will call b a *microlocal square root* of a.

Similar construction: if $a(\mathbf{x}, \boldsymbol{\xi}) \neq 0$ in $\mathcal{P}$, then there is $c(\mathbf{x}, D)$ of order $-m$ so that

$$c(\mathbf{x}, D)a(\mathbf{x}, D)u - u,\ a(\mathbf{x}, D)c(\mathbf{x}, D)u - u \in \mathcal{E}(\mathbf{R}^n)$$

for any $u \in \mathcal{E}'(\mathbf{R}^n)$ with $WF(u) \subset \mathcal{P}$.

Moreover, $c_{-m}(\mathbf{x}, \boldsymbol{\xi}) = 1/a_m(\mathbf{x}, \boldsymbol{\xi}),\ (\mathbf{x}, \boldsymbol{\xi}) \in \mathcal{P}$. Will call b a *microlocal inverse* of a.

Application: the Square Root Operator

$$a(x, z, D_t, D_x) = \frac{\partial^2}{\partial x^2} - \frac{4}{v(x,z)^2}\frac{\partial^2}{\partial t^2} = \frac{4}{v(x,z)^2} D_t^2 - D_x^2$$

is

$$a(x, z, \tau, \xi) = \frac{4}{v(x,z)^2}\tau^2 - \xi^2$$

For $\delta > 0$, set

$$\mathcal{P}_\delta(z) = \left\{ (x, t, \xi, \tau) : \frac{4}{v(x,z)^2}\tau^2 > (1+\delta)\xi^2 \right\}$$

The SSR Operator

Then according to the last slide, there is an order 1 ΨDO-valued function of z, $b(x,z,D_t,D_x)$, with principal symbol

$$b_1(x,z,\tau,\xi) = \sqrt{\frac{4}{v(x,z)^2}\tau^2 - \xi^2} = \tau\sqrt{\frac{4}{v(x,z)^2} - \frac{\xi^2}{\tau^2}},\ (x,t,\xi,\tau) \in \mathcal{P}_\delta(z)$$

for which $a(x,z,D_t,D_x)u \simeq b(x,z,D_t,D_x)b(x,z,D_t,D_x)u$ if $WF(u) \subset \mathcal{P}_\delta(z)$.

b is the world-famous **single square root** ("SSR") operator - see Claerbout, IEI.

The SSR Assumption

To what extent has this construction factored the wave operator:

$$\left(\frac{\partial}{\partial z} - ib(x,z,D_x,D_t)\right)\left(\frac{\partial}{\partial z} + ib(x,z,D_x,D_t)\right)$$

$$= \frac{\partial^2}{\partial z^2} + b(x,z,D_x,D_t)b(x,z,D_x,D_t) + \frac{\partial b}{\partial z}(x,z,D_x,D_t)$$

SSR Assumption: For some $\delta > 0$, the wavefield w satisfies

$$(x,z,t,\xi,\zeta,\tau) \in WF(w) \;\Rightarrow\; (x,t,\xi,\tau) \in \mathcal{P}_\delta(z) \text{ and } \zeta\tau > 0$$

This statement has a ray-theoretic interpretation (which will eventually make sense): rays carrying significant energy are nowhere horizontal. Along any such ray, z decreases as t increases - *coming up!*

$$\tilde{w}(x,z,t) = \left(\frac{\partial}{\partial z} + ib(x,z,D_x,D_t)\right) w(x,z,t)$$

$$b(x,z,D_x,D_t)b(x,z,D_x,D_t)w \simeq \left(\frac{4}{v(x,z)^2}D_t^2 - D_x^2\right) w$$

with a smooth error, so

$$\left(\frac{\partial}{\partial z} - ib(x,z,D_x,D_t)\right) \tilde{w}(x,z,t) = -\frac{2r(x,z)}{v(x,z)^2}\delta(t)$$

$$+i\left(\frac{\partial}{\partial z}b(x,z,D_x,D_t)\right) w(x,z,t)$$

(since b depends on z, the z deriv. does not commute with b). So $\tilde{w} = \tilde{w}_0 + \tilde{w}_1$, where

$$\left(\frac{\partial}{\partial z} - ib(x,z,D_x,D_t)\right)\tilde{w}_0(x,z,t) = -\frac{2r(x,z)}{v(x,z)^2}\delta(t)$$

(this is the **SSR modeling equation**)

$$\left(\frac{\partial}{\partial z} - ib(x,z,D_x,D_t)\right)\tilde{w}_1(x,z,t) = i\left(\frac{\partial}{\partial z}b(x,z,D_x,D_t)\right)w(x,z,t)$$

Claim: $WF(\tilde{w}_1) \subset WF(w)$. Granted this $\Rightarrow WF(\tilde{w}_0) \subset WF(w)$ also.

Upshot: SSR modeling

$$\tilde{F}_0[v]r(x_s, z_s, t) = \tilde{w}_0(x_s, z_s, t)$$

produces the same singularities (i.e. the same waves) as exploding reflector modeling, so is as good a basis for migration.

SSR migration: assume that sources all lie on $z_s = 0$.

$$< \tilde{F}_0[v]^* d, r > = < d, \tilde{F}_0[v] r >$$

$$= \int dx_s \int dt\, d(x_s, t) \tilde{w}_0(x_s, 0, t)$$

$$= \int dx_s \int dt \int dz\, d(\bar{x}_s, t)\delta(z)\tilde{w}_0(x_s, z, t)$$

Define the adjoint field q by

$$\left(\frac{\partial}{\partial z} - b(x, z, D_x, D_t)\right) q(x, z, t) = d(x, t)\delta(z),\;\; q(x, z, t) \equiv 0,\; z < 0$$

which is equivalent to solving the initial value problem

$$\left(\frac{\partial}{\partial z} - ib(x, z, D_x, D_t)\right) q(x, z, t) = 0,\; z > 0;,\;\; q(x, 0, t) = d(x, t)$$

Insert in expression for inner product, integrate by parts, use self-adjointness of b, get

$$< d, \tilde{F}_0[v]r >= \int dx \int dz\, \frac{2r(x, z)}{v(x, z)^2} q(x, z, 0)$$

whence

$$\tilde{F}_0[v]^* d(x,z) = \frac{2}{v(x,z)^2} q(x,z,0)$$

Standard description of the SSR migration algorithm:

- downward continue data (i.e. solve for q)
- image at $t=0$.

The art of SSR migration: computable approximations to $b(x,z,D_x,D_t)$ - swimming pool operator, many successors.

Proof of the Claim

Unfinished business: proof of claim

Depends on celebrated **Propagation of Singularities** theorem of Hörmander (1970).

Given symbol $p(\mathbf{x}, \boldsymbol{\xi})$, order m, with asymptotic expansion, define *bicharateristics* as solutions $(\mathbf{x}(t), \boldsymbol{\xi}(t))$ of Hamiltonian system

$$\frac{d\mathbf{x}}{dt} = \frac{\partial p}{\partial \boldsymbol{\xi}}(\mathbf{x}, \boldsymbol{\xi}), \;\; \frac{d\boldsymbol{\xi}}{dt} = -\frac{\partial p}{\partial \mathbf{x}}(\mathbf{x}, \boldsymbol{\xi})$$

with $p(\mathbf{x}(t), \boldsymbol{\xi}(t)) \equiv 0$.

Theorem: Suppose $p(\mathbf{x}, D)u = f$, and suppose that for $t_0 \leq t \leq t_1$, $(\mathbf{x}(t), \boldsymbol{\xi}(t)) \notin WF(f)$. Then either $\{(\mathbf{x}(t), \boldsymbol{\xi}(t)) : t_0 \leq t \leq t_1\} \subset WF(u)$ or $\{(\mathbf{x}(t), \boldsymbol{\xi}(t)) : t_0 \leq t \leq t_1\} \subset T^*(\mathbf{R}^n) - WF(u)$.

P of S has at least two distinct proofs:

- Nirenberg, 1972
- Hörmander, 1970 (in Taylor, 1981)

Proof of claim: check that bicharacteristics for SSR operator are just upcoming rays of geom. optics for wave equation. These pass into $t < 0$ where RHS is smooth, also initial condn at large z is smooth - so each ray has one "end" outside of $WF(\tilde{w}_1)$. If ray carries singularity, must pass of WF of w, but then it's entirely contained by P of S applied to w. **q. e. d.**

Reverse Depth Migration, Prestack

Nonzero offset ("prestack"): starting point is integral representation of the scattered field

$$F[v]r(\mathbf{x}_r, t; \mathbf{x}_s) = \frac{\partial^2}{\partial t^2} \int dx \, \frac{2r(\mathbf{x})}{v(\mathbf{x})^2} \int ds \, G(\mathbf{x}_r, t - s; \mathbf{x}) G(\mathbf{x}_s, s; \mathbf{x})$$

By analogy with zero offset case, would like to view this as "exploding reflectors in both directions": reflectors propagate energy upward to sources and to receivers.

However can't do this because reflection location is *same* for both.

The "survey sinking" idea

Bold stroke: introduce a new space variable $\mathbf{y}$ (a "sunken source", think of $\mathbf{x}$ as a "sunken receiver"), define

$$\tilde{F}[v]R(\mathbf{x}_r, t; \mathbf{x}_s) = \frac{\partial^2}{\partial t^2} \int \int dx\, dy\, R(\mathbf{x}, \mathbf{y}) \int ds\, G(\mathbf{x}_r, t - s; \mathbf{x}) G(\mathbf{x}_s, s; \mathbf{y})$$

and note that $\tilde{F}[v]R = F[v]r$ if

$$R(\mathbf{x}, \mathbf{y}) = \frac{2r}{v^2} \left(\frac{\mathbf{x} + \mathbf{y}}{2}\right) \delta(\mathbf{x} - \mathbf{y})$$

This trick decomposes $F[v]$ into *two* "exploding reflectors":

$$\tilde{F}[v]R(\mathbf{x}_r, t; \mathbf{x}_s) = u(\mathbf{x}, t; \mathbf{x}_s)|_{\mathbf{x}=\mathbf{x}_r}$$

where

$$\left(\frac{1}{v(\mathbf{x})^2}\frac{\partial^2}{\partial t^2} - \nabla_{\mathbf{x}}^2\right) u(\mathbf{x}, t; \mathbf{x}_s) = \int dy\, R(\mathbf{x}, \mathbf{y}) G(\mathbf{x}_s, t; \mathbf{y})$$

$$\equiv w_s(\mathbf{x}_s, t; \mathbf{x})$$

("upward continue the receivers"),

$$\left(\frac{1}{v(\mathbf{y})^2}\frac{\partial^2}{\partial t^2} - \nabla_{\mathbf{y}}^2\right) w_s(\mathbf{y}, t; \mathbf{x}) = R(\mathbf{x}, \mathbf{y})\delta(t)$$

("upward continue the sources").

This factorization of $F[v]$ ($r \mapsto R \mapsto \tilde{F}[v]R$) leads to a reverse time computation of adjoint $\tilde{F}[v]^*$ - will discuss this later.

It's equally possible to continue the receivers first, then the sources, which leads to

$$\left(\frac{1}{v(\mathbf{y})^2}\frac{\partial^2}{\partial t^2} - \nabla^2_{\mathbf{y}}\right) u(\mathbf{x}_r, t; \mathbf{y}) = \int dx\, R(\mathbf{x}, \mathbf{y}) G(\mathbf{x}_r, t; \mathbf{x})$$

$$\equiv w_r(\mathbf{x}_r, t; \mathbf{y})$$

("upward continue the sources"),

$$\left(\frac{1}{v(\mathbf{x})^2}\frac{\partial^2}{\partial t^2} - \nabla^2_{\mathbf{x}}\right) w_r(\mathbf{x}, t; \mathbf{y}) = R(\mathbf{x}, \mathbf{y})\delta(t)$$

("upward continue the receivers").

The DSR Assumption

Apply reverse depth concept: as before, go 2D temporarily, $\mathbf{x} = (x, z_r)$, $\mathbf{y} = (y, z_s)$, all sources and receivers on $z = 0$.

Double Square Root ("DSR") assumption: For some $\delta > 0$, the wavefield u satisfies

$$(x, z_r, t, y, z_s, \xi, \zeta_s, \tau, \eta, \zeta_r) \in WF(u) \;\Rightarrow$$

$$(x, t, \xi, \tau) \in \mathcal{P}_\delta(z_r),\;\; (y, t, \eta, \tau) \in \mathcal{P}_\delta(z_s),\;\; \text{and}\;\; \zeta_r \tau > 0,\;\; \zeta_s \tau > 0,$$

As for SSR, there is a ray-theoretic interpretation: rays from source and receiver to scattering point stay away from the vertical and decrease in z for increasing t, i.e. they are all upcoming.

Since z will be singled out (and eventually $R(\mathbf{x},\mathbf{y})$ will have a factor of $\delta(\mathbf{x},\mathbf{y})$), impose the constraint that

$$R(x,z,x,z_s) = \tilde{R}(x,y,z)\delta(z-z_s)$$

Define upcoming projections as for SSR:

$$\tilde{w}_s = \left(\frac{\partial}{\partial z_s} + ib(y,z_s,D_y,D_t)\right) w_s,$$

$$\tilde{w}_r = \left(\frac{\partial}{\partial z_r} + ib(x,z_r,D_x,D_t)\right) w_r,$$

$$\tilde{u} = \left(\frac{\partial}{\partial z_s} + ib(y,z_s,D_y,D_t)\right)\left(\frac{\partial}{\partial z_r} + ib(x,z_r,D_x,D_t)\right) u$$

Except for lower order commutators which we justify throwing away as before,

$$\left(\frac{\partial}{\partial z_s} - ib(y, z_s, D_y, D_t)\right) \tilde{w}_s = \tilde{R}\delta(z_r - z_s)\delta(t),$$

$$\left(\frac{\partial}{\partial z_r} - ib(x, z_r, D_x, D_t)\right) \tilde{w}_r = \tilde{R}\delta(z_r - z_s)\delta(t),$$

$$\left(\frac{\partial}{\partial z_r} - ib(x, z_r, D_x, D_t)\right) \tilde{u} = \tilde{w}_s$$

$$\left(\frac{\partial}{\partial z_s} - ib(y, z_s, D_y, D_t)\right) \tilde{u} = \tilde{w}_r$$

Initial (final) conditions are that $\tilde{w}_r, \tilde{w}_s$, and $\tilde{u}$ all vanish for large z - the equations are to be solve in decreasing z (“upward continuation”).

Simultaneous upward continuation:

$$\frac{\partial}{\partial z}\tilde{u}(x,z,t;y,z) = \frac{\partial}{\partial z_r}\tilde{u}(x,z_r,t;y,z)|_{z=z_r} + \frac{\partial}{\partial z_r}\tilde{u}(x,z,t;y,z_s)|_{z=z_s}$$

$$= [ib(x,z_r,D_x,D_t)\tilde{u} + \tilde{w}_s + ib(y,z_s,D_y,D_t)\tilde{u} + \tilde{w}_r]_{z_r=z_s=z}$$

Since $\tilde{w}_s(y,z,t;x,z) = \tilde{w}_r(x,z,t;y,z) = \tilde{R}(x,y,z)\delta(t)$, $\tilde{u}$ is seen to satisfy the

DSR modeling equation:

$$\left(\frac{\partial}{\partial z} - ib(x,z,D_x,D_t) - ib(y,z,D_y,D_t)\right)\tilde{u}(x,z,t;y,z) = 2\tilde{R}(x,y,z)\delta(t)$$

$$\tilde{F}[v]\tilde{R}(x_r,t;x_s) = \tilde{u}(x_r,0,t;x_s,0)$$

DSR Migration

Computation of adjoint follows same pattern as for SSR, and leads to

DSR migration equation: solve

$$\left(\frac{\partial}{\partial z} - ib(x,z,D_x,D_t) - ib(y,z,D_y,D_t)\right)\tilde{q}(x,y,z,t) = 0$$

in *increasing* z with initial condition at $z = 0$:

$$\tilde{q}(x_r,x_s,0,t) = d(x_r,x_s,t)$$

Then $\tilde{F}[v]^*d(x,y,z) = \tilde{q}(x,y,z,0)$

The physical DSR model has $\tilde{R}(x,y,z) = r(x,z)\delta(x-y)$, so final step in DSR computation of $F[v]^*$ is adjoint of $r \mapsto \tilde{R}$:

$$F[v]^*d(x,z) = \tilde{q}(x,x,z,0)$$

Standard description of DSR migration

(See Claerbout, IEI):

- downward continue sources and receivers (solve DSR migration equation)
- image at $t = 0$ and zero offset ($x = y$)

Another moniker: "survey sinking": DSR field $\tilde{q}$ is (related to) the field that you would get by conducting the survey with sources and receivers at depth z. At any given depth, the zero-offset, time-zero part of the field is the instantaneous response to scatterers on which source = receiver is sitting, therefore constitutes an image.

As for SSR, the art of DSR migration is in the approximation of the DSR operator.

Remarks

Stolk and deHoop (2001) derived DSR modeling and migration via a more systematic argument than that used here, involving ΨDO matrix factorization of the wave equation written as a first order evolution system in z. This idea goes back to Taylor (1975) who used it to show that singularities propagating along bicharacteristics reflect as expected at boundaries.

Stolk (2003) has also carried out a very careful global construction of a family of SSR ΨDOs which are of non-classical type at near-horizontal directions ("nearly evanescent waves"). This construction should lead to more reliable discretizations.

The last part of the course will present the various apparently ad-hoc "prestack modeling" ideas within a unified framework.

Mathematics of Seismic Imaging
Part III

William W. Symes

PIMS, July 2005

A step beyond linearization: velocity analysis

Velocity Analysis

Partially linearized seismic inverse problem ("velocity analysis"): given observed seismic data d, find smooth *velocity* $v \in \mathcal{E}(X), X \subset \mathbf{R}^3$ oscillatory *reflectivity* $r \in \mathcal{E}'(X)$ so that

$$F[v]r \simeq d$$

Acoustic partially linearized model: acoustic potential field u and its perturbation δu solve

$$\left(\frac{1}{v^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) u = \delta(t)\delta(\mathbf{x} - \mathbf{x}_s), \;\; \left(\frac{1}{v^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) \delta u = 2r\nabla^2 u$$

plus suitable bdry and initial conditions.

$$F[v]r = \left.\frac{\partial \delta u}{\partial t}\right|_Y$$

data acquisition manifold $Y = \{(\mathbf{x}_r, t; \mathbf{x}_s)\} \subset \mathbf{R}^7$, dimn $Y \leq 5$ (many idealizations here!).

$F[v] : \mathcal{E}'(X) \to \mathcal{D}'(Y)$ is a linear map (FIO of order 1), but dependence on v is quite nonlinear, so this inverse problem is nonlinear.

Agenda:

- reformulation of inverse problem via *extensions*
- "standard processing" extension and standard VA
- the surface oriented extension and standard MVA
- the ΨDO property and why it's important
- global failure of the ΨDO property for the SOE
- Claerbout's depth oriented extension has the ΨDO property
- differential semblance

Extensions

Extension of $F[v]$: manifold $\bar{X}$ and maps $\chi : \mathcal{E}'(X) \to \mathcal{E}'(\bar{X})$, $\bar{F}[v] : \mathcal{E}'(\bar{X}) \to \mathcal{D}'(Y)$ so that

$$\begin{array}{ccccc} & & \bar{F}[v] & & \\ & \mathcal{E}'(\bar{X}) & \longrightarrow & \mathcal{D}'(Y) & \\ \chi & \uparrow & & \uparrow & \mathrm{id} \\ & \mathcal{E}'(X) & \longrightarrow & \mathcal{D}'(Y) & \\ & & F[v] & & \end{array}$$

commutes.

Invertible extension: $\bar{F}[v]$ has a *right parametrix* $\bar{G}[v]$, i.e. $I - \bar{F}[v]\bar{G}[v]$ is smoothing. [The trivial extension - $\bar{X} = X, \bar{F} = F$ - is virtually never invertible.] Also χ has a *left inverse* η.

Reformulation of inverse problem: given d, find v so that $\bar{G}[v]d \in \mathcal{R}(\chi)$ (implicitly determines r also!).

Reformulation of inverse problem

Given d, find v so that $\bar{G}[v]d \in$ the range of χ.

Claim: if v is so chosen, then $[v,r]$ solves partially linearized inverse problem with $r = \eta\bar{G}[v]d$.

Proof: Hypothesis means

$$\bar{G}[v]d = \chi r$$

for some r (whence necessarily $r = \eta\bar{G}[v]d$), so

$$d \simeq \bar{F}[v]\bar{G}[v]d = \bar{F}[v]\chi r = F[v]r$$

Q. E. D.

Example 1: Standard VA extension

Treat each CMP *as if it were the result of an experiment performed over a layered medium*, but permit the layers to vary with midpoint.

Thus $v = v(z), r = r(z)$ for purposes of analysis, but at the end $v = v(\mathbf{x}_m, z), r = r(\mathbf{x}_m, z)$.

$$F[v]R(\mathbf{x}_m, h, t) \simeq A(\mathbf{x}_m, h, z(\mathbf{x}_m, h, t))R(\mathbf{x}_m, z(\mathbf{x}_m, h, t))$$

Here $z(\mathbf{x}_m, h, t)$ is the inverse of the 2-way traveltime

$$t(\mathbf{x}_m, h, z) = 2\tau(\mathbf{x}_m + (h, 0, z), \mathbf{x}_m)_{v=v(\mathbf{x}_m,z)}$$

computed with the layered velocity $v(\mathbf{x}_m, z)$, i.e.
$z(\mathbf{x}_m, h, t(\mathbf{x}_m, h, z')) = z'$.

That is, $F[v]$ *is a change of variable followed by multiplication by a smooth function*. NB: industry standard practice is to use vertical traveltime t_0 instead of z for depth variable.

Can write this as $F[v] = \bar{F}[v]S^*$, where $\bar{F}[v] = N[v]^{-1}M[v]^{-1}$ has right parametrix $\bar{G}[v] = M[v]N[v]$:

$N[v]$ = **NMO operator** $N[v]d(\mathbf{x}_m, h, z) = d(\mathbf{x}_m, h, t(\mathbf{x}_m, h, z))$

$M[v]$ = multiplication by A

S = **stacking operator**

$$Sf(\mathbf{x}_m, z) = \int dh\, f(\mathbf{x}_m, h, z), \;\; S^*r(\mathbf{x}_m, h, z) = r(\mathbf{x}, z)$$

Identify as extension: $\bar{F}[v], \bar{G}[v]$ as above, $X = \{\mathbf{x}_m, z\}, H = \{h\}, \bar{X} = X \times H, \chi = S^*, \eta = S$ - the invertible extension properties are clear.

Standard names for the Standard VA extension objects: $\bar{F}[v]$ = "inverse NMO", $\bar{G}[v]$ = "NMO" [often the multiplication op $M[v]$ is neglected]; η = "stack", χ = "spread"

How this is used for velocity analysis: Look for v that makes $\bar{G}[v]d \in \mathcal{R}(\chi)$

So what is $\mathcal{R}(\chi)$? $\chi[r](\mathbf{x}_m, z, h) = r(\mathbf{x}_m, z)$ Anything in range of χ is *independent of* h. Practical issues $\Rightarrow$ replace "independent of" with "smooth in".

Flatten them gathers!

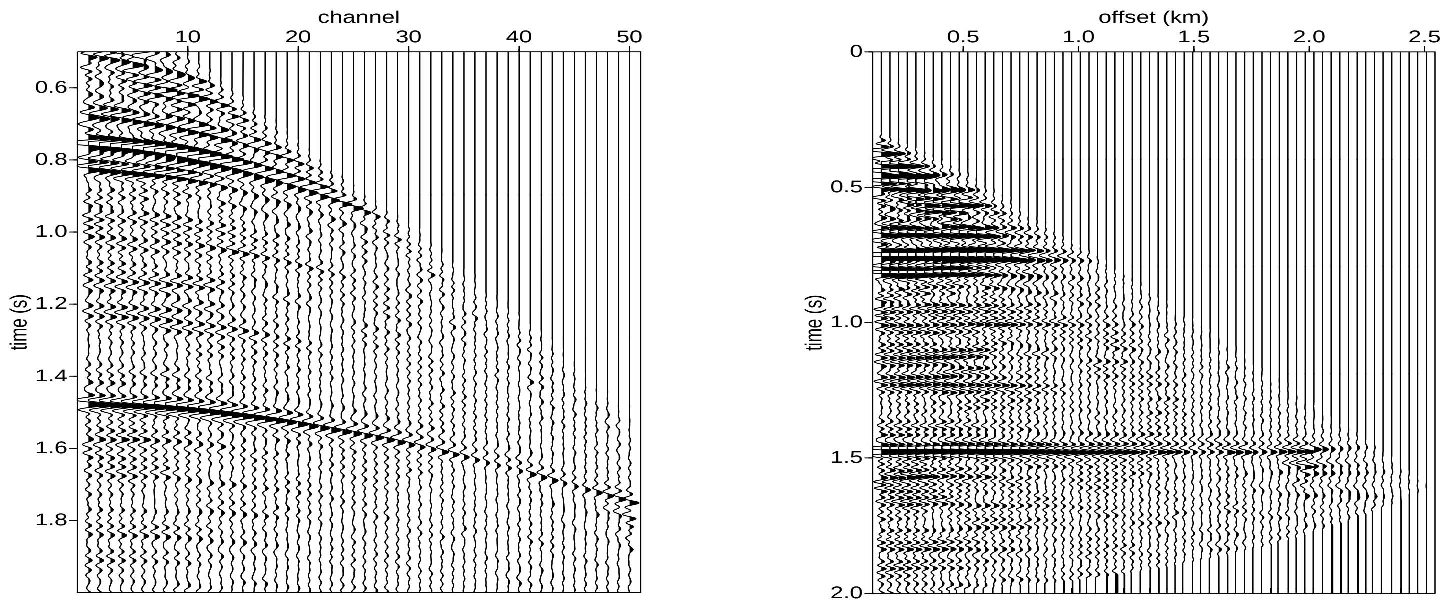
Inverse problem reduced to: adjust v to make $\bar{G}[v]d^{\mathrm{obs}}$ smooth in h, i.e. *flat* in z, h display for each $\mathbf{x}_m$ (*NMO-corrected CMP*).

Replace z with t_0, v with v_{RMS} em localizes computation: reflection through $\mathbf{x}_m, t_0, 0$ *flattened* by adjusting $v_{\mathrm{RMS}}(\mathbf{x}_m, t_0) \Rightarrow$ 1D search, do by visual inspection.

Various aids - NMO corrected CMP gathers, velocity spectra, etc.

See: Claerbout: *Imaging the Earth's Interior*

WWS: MGSS 2000 notes



,

Left: part of survey (d) from North Sea (thanks: Shell Research), lightly preprocessed.
Right: restriction of $\bar{G}[v]d$ to $\mathbf{x}_m = \text{const}$ (function of depth, offset): shows rel. sm'ness in h (offset) for properly chosen v.

Example 2: Surface oriented or standard MVA extension

. Standard VA only works where Earth is "nearly layered". Where this fails, replace NMO by prestack migration.

Version based on common offset modeling/migration: Σ_h = set of half-offsets in data, $\bar{X} = X \times \Sigma_h$, $\chi[r](\mathbf{x},\mathbf{h}) = r(\mathbf{x})$.

$$\bar{F}[v]\bar{r}(\mathbf{x}_s,t,\mathbf{h}) = \frac{\partial^2}{\partial t^2}\int dx\,\bar{r}(\mathbf{x},\mathbf{h})\int ds\,G(\mathbf{x}_s+2\mathbf{h},t-s;\mathbf{x})G(\mathbf{x}_s,s;\mathbf{x})$$

Note that this operator is "block diagonal" in $\mathbf{h}$.

Properties of SOE

Beylkin (1985), Rakesh (1988): if $\|v\|_{C^2(X)}$ "not too big", then

- $\bar{F}$ has **the** Ψ**DO property**: $\bar{F}\bar{F}^*$ is ΨDO
- singularities of $\bar{F}\bar{F}^*d \subset$ singularities of d
- straightforward construction of right parametrix $\bar{G} = \bar{F}^*Q$, Q = ΨDO, also as generalized Radon Transform - explicitly computable.

Range of χ (offset version): $\bar{r}(\mathbf{x}, \mathbf{h})$ independent of $\mathbf{h}$ $\Rightarrow$ "semblance principle": find v so that $\bar{G}[v]d^{\mathrm{obs}}$ is independent of $\mathbf{h}$. Practical limitations $\Rightarrow$ replace "independent of $\mathbf{h}$" by "smooth in $\mathbf{h}$".

Industrial MVA

Application of these ideas = industrial practice of migration velocity analysis.

Idea: twiddle v until $\bar{G}[v]d^{\rm obs}$ is smooth in $\mathbf{h}$.

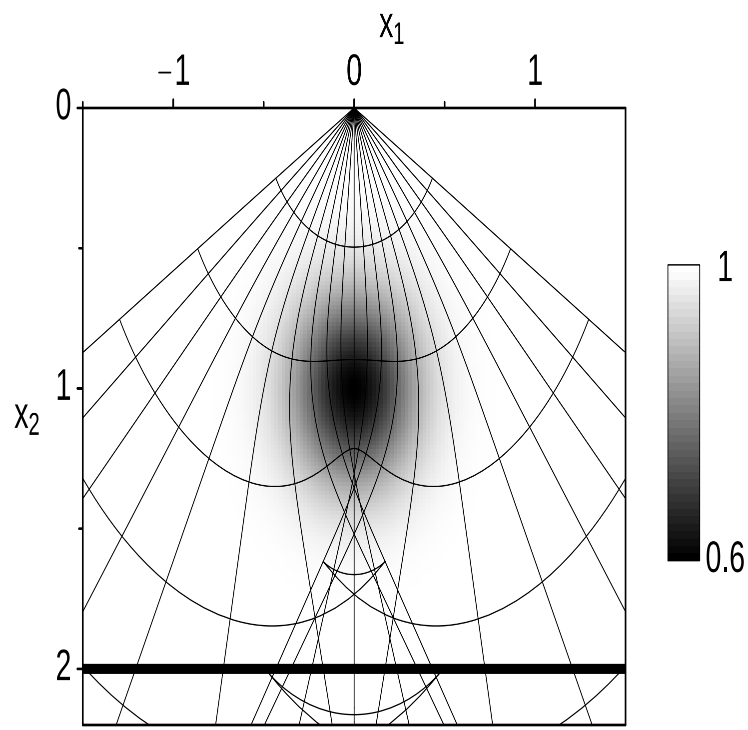
Since it is hard to inspect $\bar{G}[v]d^{\rm obs}(x,y,z,h)$, pull out subset for constant x,y = **common image gather** ("CIG"): display function of z,h for fixed x,y. These play same role as NMO corrected CMP gathers in layered case.

Try to adjust v so that selected CIGs are *flat* - just as in Standard VA. This is much harder, as there is no RMS velocity trick to localize the computation - each CIG depends globally on v.

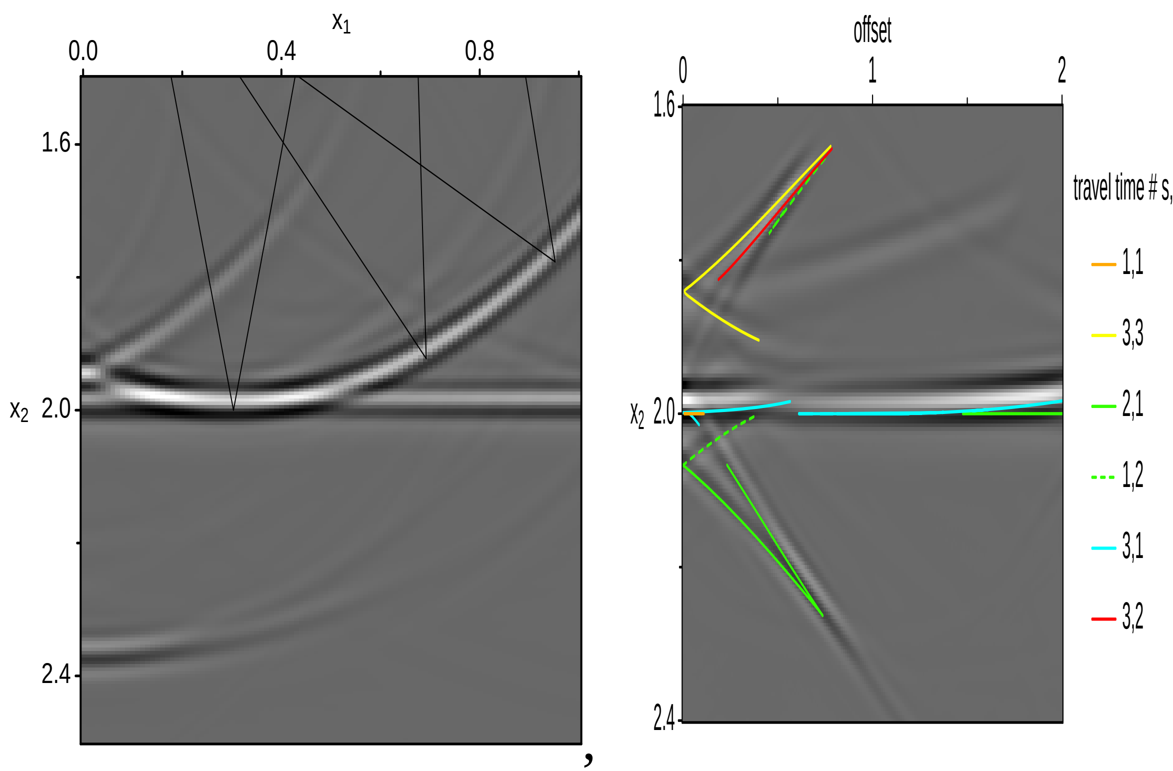
Description, some examples: Yilmaz, *Seismic Data Processing*.

Bad news

Nolan (1997), Stolk & WWS (2004): big trouble! In general, standard extension does **not** have the ΨDO property. Geometric optics analysis: for $\|v\|_{C^2(X)}$ "large", multiple rays connect source, receiver to reflecting points in X; block diagonal structure of $\bar{F}[v]$ $\Rightarrow$ info necessary to distinguish multiple rays is *projected out*.



Example (Stolk & WWS, 2001): Gaussian lens over flat reflector at depth z ($r(\mathbf{x}) = \delta(x_1 - z)$, $x_1 =$ depth).



Left: Const. h slice of $\bar{G}d$: several refl. points corresponding to same singularity in d^{obs}.
Right: CIG (const. x, y slice) of $\bar{G}d$: not smooth in h!

Example 3: Claerbout's depth oriented extension

Standard MVA extension only works when Earth has simple ray geometry. Claerbout (1971) proposed alternative extension:

Σ_d = somewhat arbitrary set of vectors near 0 ("offsets"), $\bar{X} = X \times \Sigma_d$, $\chi[r](\mathbf{x},\mathbf{h}) = r(\mathbf{x})\delta(\mathbf{h})$, $\eta[\bar{r}](\mathbf{x}) = \bar{r}(\mathbf{x},0)$

$$\bar{F}[v]\bar{r}(\mathbf{x}_s,t,\mathbf{x}_r) = \frac{\partial^2}{\partial t^2}\int dx \int_{\Sigma_d} dh\, \bar{r}(\mathbf{x},\mathbf{h}) \int ds\, G(\mathbf{x}_s, t-s; \mathbf{x}+2\mathbf{h})G(\mathbf{x}_r, s; \mathbf{x})$$

$$= \frac{\partial^2}{\partial t^2}\int dx \int_{\mathbf{x}+2\Sigma_d} dy\, \bar{r}(\mathbf{x},\mathbf{y}-\mathbf{x}) \int ds\, G(\mathbf{x}_s, t-s; \mathbf{y})G(\mathbf{x}_r, s; \mathbf{x})$$

NB: in this formulation, there appears to be too many model parameters.

Shot record modeling

for each $\mathbf{x}_s$ solve

$$\bar{F}[v]\bar{r}(\mathbf{x}_r, t; \mathbf{x}_s) = u(\mathbf{x}, t; \mathbf{x}_s)|_{\mathbf{x}=\mathbf{x}_r}$$

where

$$\left(\frac{1}{v(\mathbf{x})^2}\frac{\partial^2}{\partial t^2} - \nabla^2_{\mathbf{x}}\right) u(\mathbf{x}, t; \mathbf{x}_s) = \int_{\mathbf{x}+2\Sigma_d} dy\, \bar{r}(\mathbf{x}, \mathbf{y}) G(\mathbf{y}, t; \mathbf{x}_s)$$

$$\left(\frac{1}{v(\mathbf{y})^2}\frac{\partial^2}{\partial t^2} - \nabla^2_{\mathbf{y}}\right) G(\mathbf{y}, t; \mathbf{x}_s) = \delta(t)\delta(\mathbf{x}_s - \mathbf{y})$$

Finite difference scheme: form RHS for eqn 1, step u, G forward in t.

Computing $\bar{G}[v]$

Instead of parametrix, be satisfied with adjoint.

Reverse time adjoint computation - specify adjoint field as in standard reverse time prestack migration:

$$\left(\frac{1}{v(\mathbf{x})^2}\frac{\partial^2}{\partial t^2} - \nabla^2_{\mathbf{x}}\right) w(\mathbf{x},t;\mathbf{x}_s) = \int dx_r\, d(\mathbf{x}_r,t;\mathbf{x}_s)\delta(\mathbf{x}-\mathbf{x}_r)$$

with $w(\mathbf{x},t;\mathbf{x}_s) = 0, t >> 0$. Then

$$\bar{F}[v]^* d(\mathbf{x},\mathbf{h}) = \int dx_s \int dt\, G(\mathbf{x}+2\mathbf{h},t;\mathbf{x}_s) w(\mathbf{x},t;\mathbf{x}_s)$$

i.e. exactly the same computation as for reverse time prestack, except that crosscorrelation occurs at an offset $2\mathbf{h}$.

Nomenclature

NB: the "usual computation" of $\bar{G}[v]$ is either DSR or a variant of shot record computation of previous slide using depth extrapolation. $\mathbf{h}$ is usually restricted to be horizontal, i.e. $h_3 = 0$.

Common names: shot-geophone or survey-sinking migration (with DSR), or shot record migration.

"Downward continue sources and receivers, image at $t = 0, h = 0$"

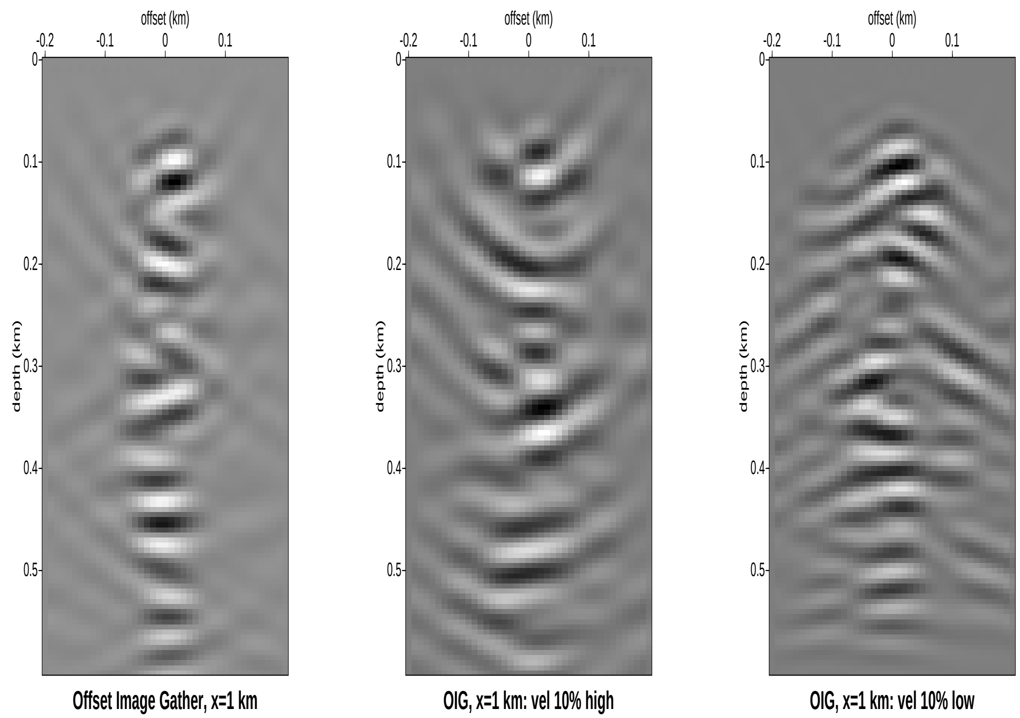
These are what is typically meant by "wave equation migration"!

What should be the character of the image when the velocity is correct?

Hint: for simulation of seismograms, the input reflectivity had the form $r(\mathbf{x})\delta(\mathbf{h})$.

Therefore guess that when velocity is correct, *image is concentrated near* $h = 0$.

Examples: 2D finite difference implementation of reverse time method. Correct velocity $\equiv 1$. Input reflectivity used to generate synthetic data: random! For output reflectivity (image of $\bar{F}[v]^*$), constrain offset to be horizontal: $\bar{r}(\mathbf{x}, \mathbf{h}) = \tilde{r}(\mathbf{x}, h_1)\delta(h_3)$. Display CIGs (i.e. x_1 =const. slices).



, ,

Two way reverse time horizontal offset S-G image gathers of data from random reflectivity, constant velocity. From left to right: correct velocity, 10% high, 10% low.

Stolk and deHoop, 2001

Claerbout extension has the ΨDO property, at least when restricted to $\bar{r}$ of the form $\bar{r}(\mathbf{x},\mathbf{h}) = R(\mathbf{x},h_1,h_2)\delta(h_3)$, and under DSR assumption.

Sketch of proof (after Rakesh, 1988):

This will follow from *injectivity* of wavefront or *canonical relation* $C_{\bar{F}} \subset T^*(\bar{X}) - \{\mathbf{0}\} \times T^*(Y) - \{\mathbf{0}\}$ which describes singularity mapping properties of $\bar{F}$:

$$(\mathbf{x},\mathbf{h},\xi,\nu,\mathbf{y},\eta) \in C_{F_{\delta}[v]} \Leftrightarrow$$

$$\text{for some } u \in \mathcal{E}'(\bar{X}),\ (\mathbf{x},\mathbf{h},\xi,\nu) \in WF(u),\ \text{and } (\mathbf{y},\eta) \in WF(\bar{F}u)$$

Characterization of $C_{\bar{F}}$

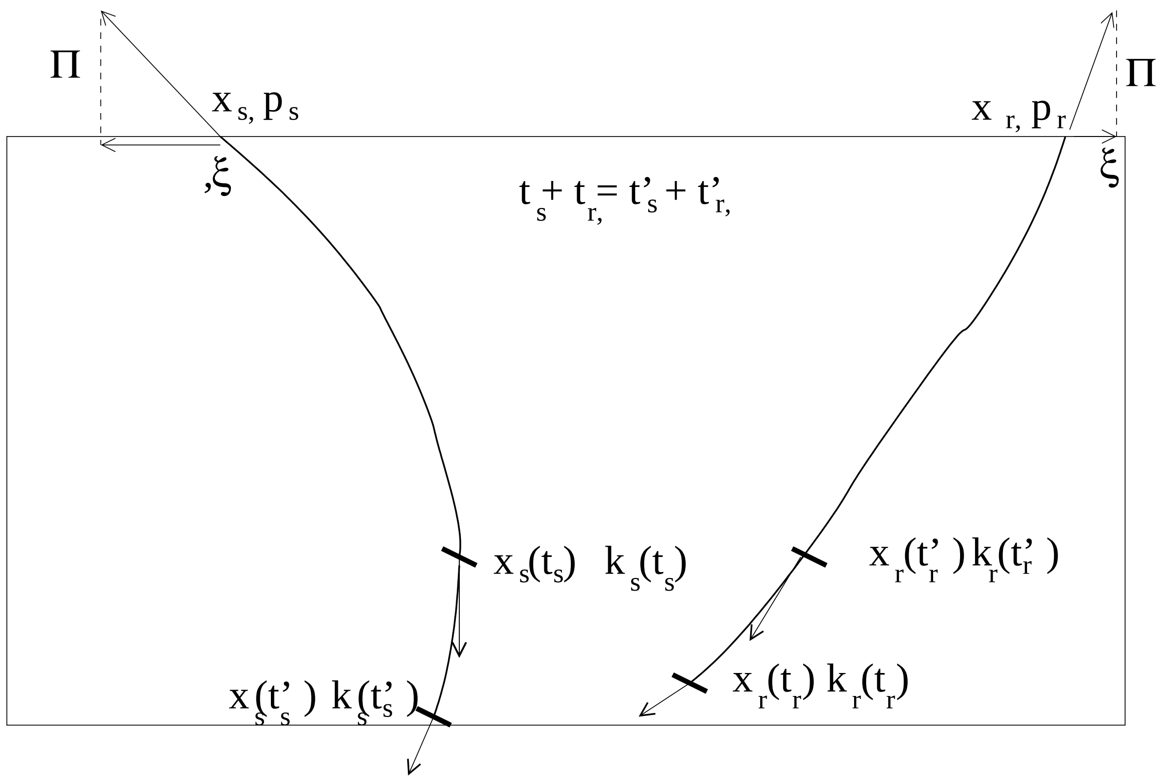
$$((\mathbf{x},\mathbf{h},\xi,\nu),(\mathbf{x}_s,t,\mathbf{x}_r,\xi_{\mathbf{s}},\tau,\xi_{\mathbf{r}}))\in C_{\bar{F}}[v]\subset T^*(\bar{X})-\{\mathbf{0}\}\times T^*(Y)-\{\mathbf{0}\}$$

$\Leftrightarrow$ there are *rays of geometric optics* $(\mathbf{X}_s,\mathbf{\Xi}_s)$**,** $(\mathbf{X}_r,\mathbf{\Xi}_r)$ and times t_s,t_r so that

$$\Pi(\mathbf{X}_s(0),t,\mathbf{X}_r(0),\mathbf{\Xi}_s(0),\tau,\mathbf{\Xi}_r(0))=(\mathbf{x}_s,t,\mathbf{x}_r,\xi_s,\tau,\xi_r),$$

$$\mathbf{X}_s(t_s)=\mathbf{x},\mathbf{X}_r(t_r)=\mathbf{x}+2\mathbf{h},\ t_s+t_r=t,$$

$$\mathbf{\Xi}_s(t_s)+\mathbf{\Xi}_r(t_r)||\xi,\mathbf{\Xi}_s(t_s)-\mathbf{\Xi}_r(t_r)||\nu$$


Π
x_{s}, p_{s}
ξ
x_{r}, p_{r}
Π
ξ
$t_s + t_r = t'_s + t'_r,$
$x_s(t_s)$ $k_s(t_s)$
$x_r(t'_r)\, k_r(t'_r)$
$x_s(t'_s)\, k_s(t'_s)$
$x_r(t_r)\, k_r(t_r)$

Proof

Uses wave equations for u, G and

- Gabor calculus: computes wave front sets of products, pullbacks, integrals, etc. See Duistermaat, Ch. 1.
- Propagation of Singularities Theorem

and that's all! [No integral representations, phase functions,...]

Note intrinsic ambiguity: if you have a ray pair, move times t_s, t_r resp. t'_s, t'_r, for which $t_s + t_r = t'_s + t'_r = t$ then you can construct two points $(\mathbf{x}, \mathbf{h}, \boldsymbol{\xi}, \nu), (\mathbf{x}', \mathbf{h}', \boldsymbol{\xi}', \nu')$ which are candidates for membership in $WF(\bar{r})$ and which satisfy the above relations with the same point in the cotangent bundle of $T^*(Y)$.

No wonder - there are too many model parameters!

Stolk and deHoop fix this ambiguity by imposing two constraints:

- DSR assumption: all rays carrying significant reflected energy (source or receiver) are upcoming.
- Restrict $\bar{F}$ to the domain $\mathcal{Z} \subset \mathcal{E}'(\bar{X})$

$$\bar{r} \in \mathcal{Z} \;\Leftrightarrow\; \bar{r}(\mathbf{x}, \mathbf{h}) = R(\mathbf{x}, h_1, h_2)\delta(h_3)$$

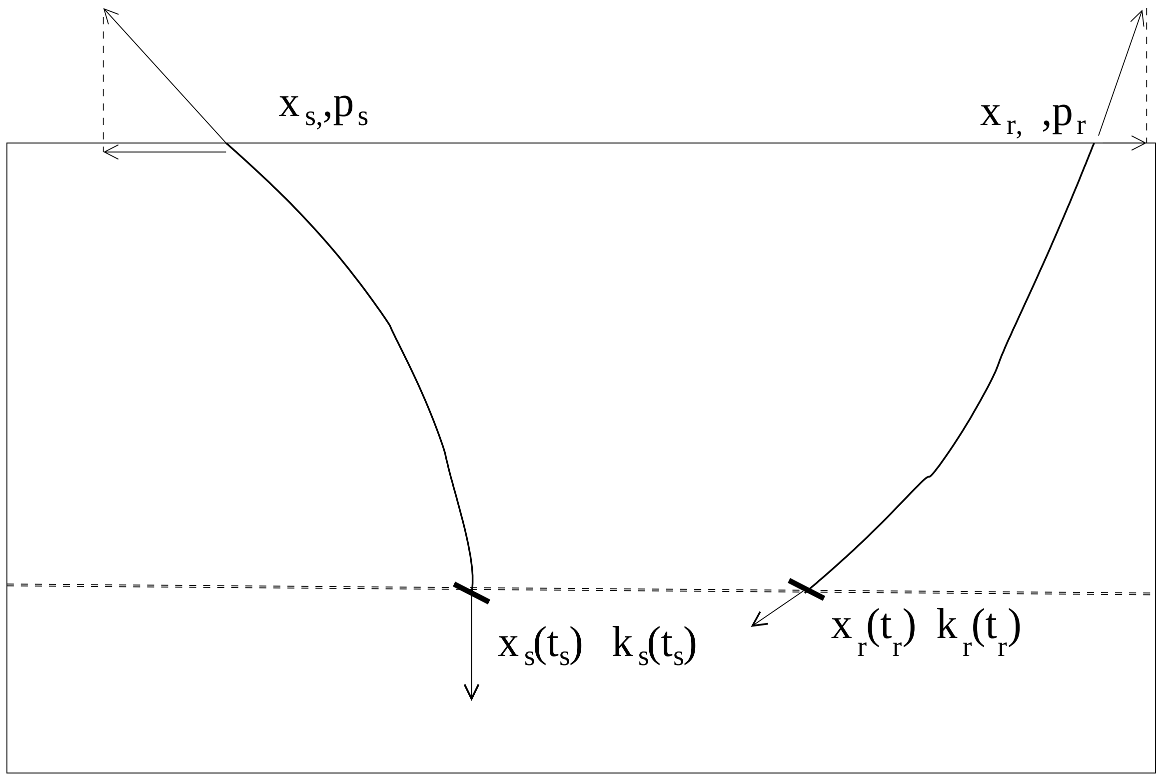
If $\bar{r} \in \mathcal{Z}$, then $(\mathbf{x}, \mathbf{h}, \boldsymbol{\xi}, \nu) \in WF(\bar{r}) \Rightarrow h_3 = 0$. So source and receiver rays in $C_{\bar{F}}$ must terminate at same depth, to hit such a point.

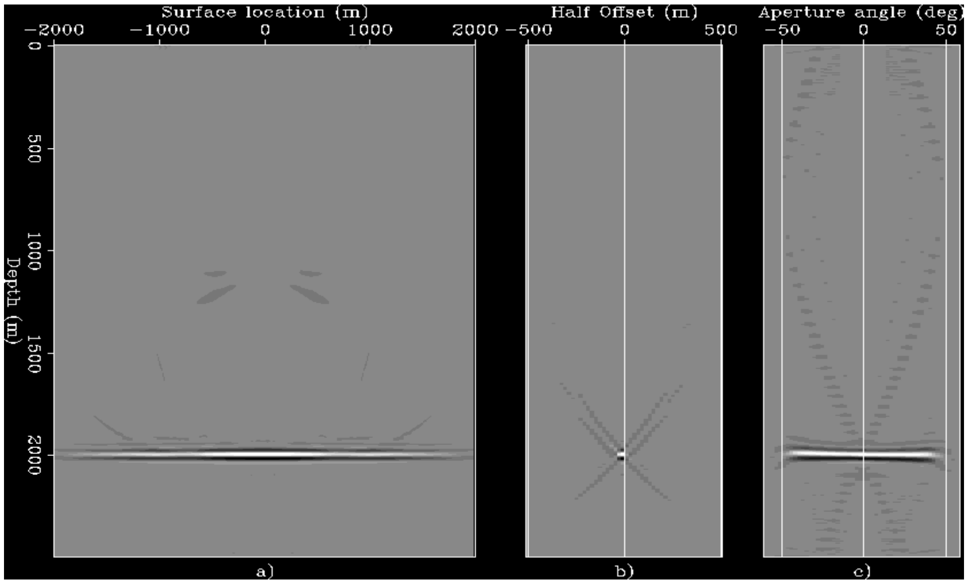
Because of DSR assumption, this fixes the traveltimes t_s, t_r.

Restricted to $\mathcal{Z}$, $C_{\bar{F}}$ is injective.

$\Rightarrow C_{\bar{F}^*\bar{F}} = I$

$\Rightarrow \bar{F}^*\bar{F}$ is ΨDO when restricted to $\mathcal{Z}$.


$x_{s,}, p_s$
$x_{r,}\ , p_r$
$x_s(t_s)\ \ k_s(t_s)$
$x_r(t_r)\ \ k_r(t_r)$



Lens data, shot-geophone migration [B. Biondi, 2002]
Left: Image via DSR. Middle: $\bar{G}[v]d$ - well-focused (at $\mathbf{h} = 0$), i.e. in range of χ to extent possible. Right: Angle CIG.

Quantitative VA

Suppose $W : \mathcal{E}'(\bar{X}) \to \mathcal{D}'(Z)$ annihilates range of χ:

$$\mathcal{E}'(X) \xrightarrow{\chi} \mathcal{E}'(\bar{X}) \xrightarrow{W} \mathcal{D}'(Z) \to 0$$

and moreover W is bounded on $L^2(\bar{X})$. Then

$$J[v;d] = \frac{1}{2}\|W\bar{G}[v]d\|^2$$

minimized when $[v, \eta\bar{G}[v]d]$ solves partially linearized inverse problem.

Construction of *annihilator* of $\mathcal{R}(F[v])$ (Guillemin, 1985):

$$d \in \mathcal{R}(F[v]) \Leftrightarrow \bar{G}[v]d \in \mathcal{R}(\chi) \Leftrightarrow W\bar{G}[v]d = 0$$

Annihilators, annihilators everywhere...

For Standard Extended Model, several popular choices:

- $W = (I - \Delta)^{-\frac{1}{2}}\nabla_{\mathbf{h}}$ ("differential semblance" - WWS, 1986)
- $W = I - \frac{1}{|H|}\int\, dh$ ("stack power" - Toldi, 1985)
- $W = I - \chi F[v]^{\dagger}\bar{F}[v] \Rightarrow$ minimizing $J[v,d]$ equivalent to least squares.

For Claerbout extension, differential semblance $W = h$.

But not many are good for much...

Since *problem is huge*, only W giving rise to differentiable $v \mapsto J[v,d]$ are useful - must be able to use Newton!!! Once again, idealize $w(t) = \delta(t)$.

Theorem (Stolk & WWS, 2003): $v \mapsto J[v,d]$ smooth $\Leftrightarrow$ W pseudodifferential.

i.e. only *differential semblance* gives rise to smooth optimization problem, *uniformly in source bandwidth*.

Numerical examples using synthetic and field data: WWS et al., Chauris & Noble 2001, Mulder & tenKroode 2002. deHoop et al. 2004.

Example: NMO-based Differential Semblance

$$J[v,d] = \frac{1}{2}\left\|\frac{\partial}{\partial h}N[v]d\right\|^2$$

(recall that $N[v]$ is the NMO operator = composition with $t(z,h)$)

Theory: under some circumstances, can show that *all stationary points are global minimizers* (WWS, TRIP annual reports '99, '01).

Example uses data from North Sea survey (thanks: Shell) with light preprocessing: cutoff ("mute") and multiple suppression (predictive decon) to enhance conformance with model, low pass filter.

Minimization of J via quasi-Newton method.

Beyond Born

Nonlinear effects not included in linearized model: *multiple reflections*. Conventional approach: treat as *coherent noise*, attempt to eliminate - active area of research going back 40+ years, with recent important developments.

Why not model this "noise"?

Proposal: *nonlinear extensions* with $F[v]r$ replaced by $\mathcal{F}[c]$. Create annihilators in same way (now also nonlinear), optimize differential semblance.

Nonlinear analog of Standard Extended Model appears to be *invertible* - in fact extended nonlinear inverse problem is *underdetermined*.

Open problems: no theory. Also must determine $w(t)$ (Lailly SEG 2003).

And so on...

- Elasticity: theory of asymptotic Born inversion at smooth background in good shape (Beylkin & Burridge 1988, deHoop & Bleistein 1997). Theory of extensions, annihilators, differential semblance partially complete (Brandsberg-Dahl et al 2003).
- Anisotropy - work of deHoop (Brandsberg-Dahl et al 2003).
- Anelasticity - in the sedimentary section, $Q = 100 - 1000$, lower in gassy sediments and near surface. No mathematical results, but some numerics - Minkoff & WWS 1997, Blanch et al 1998.
- Source determination - actually always an issue. Some success in casting as an inverse problem - Minkoff & WWS 1997, Routh et al SEG 2003.
- ...

Mathematics of Seismic Imaging: Selected References

William W. Symes

July 2005

Overviews

Bleistein, N., Cohen, J. and Stockwell, J.: *The Mathematics of Seismic Imaging, Migration, and Inversion*, Springer, 2000.

Symes, W. W.: *Mathematics of Reflection Seismology*, Lecture Notes, `www.trip.caam.rice.edu`, 1995.

Imaging (1)

Cohen, J. K. and Bleistein, N.: An inverse method for determining small variations in propagation speed, *SIAM J. Appl. Math. 32*, 1977, pp. 784-799.

Beylkin, G.: Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform, *J. Math. Phys 26*, 1985, pp. 99-108.

Miller, D., Oristaglio,, M. and Beylkin, G.: A new slant on seismic imaging, *Geophysics 52*, pp. 943-964, 1987.

Rakesh: A linearized inverse problem for the wave equation, *Comm. PDE 13*, 1988, pp. 573-601.

Ten Kroode, A. P. E., Smit, D. J., and Verdel, A. R.: A microlocal analysis of migration *Wave Motion 28*, 1998, pp. 149-172.

Imaging (2)

Nolan, C. J. and Symes, W. W.: Global solution of a linearized inverse problem for the wave equation, *Comm. PDE 22*, 1997, pp. 919-952.

deHoop, M. and Bleistein, N.: Generalized Radon Transform inversions for reflectivity in anisotropic elastic media, *Inverse Problems 16*, 1998, pp. 669-690.

Stolk, C.: Microlocal analysis of a seismic linearized inverse problem, *Wave Motion 32*, 2000, pp. 267-290.

Operto, S., Xu, S. and Lambaré, G.: Can we image quantitatively complex models with rays?, *Geophysics 65*, 2001, pp. 1223-1238.

Stolk, C. and Symes, W. W.: Kinematic artifacts in prestack depth migration, *Geophysics 68*, May-June 2004.

Velocities

Minkoff, S. E. and Symes, W. W.: Full waveform inversion of marine reflection data in the plane wave domain, *Geophysics 62*, 1997, pp. 540-553.

Brandsberg-Dahl, S., De Hoop, M., and Ursin, B.: Focussing in dip and AVA compensation on scattering angle/azimuth common image gathers, *Geophysics 68*, 2003, pp. 232-254.

Stolk, C. and Symes, W. W.: Smooth objective functionals for seismic velocity inversion, *Inverse Problems 19*, 2003, pp. 73-89.

Stolk, C. and deHoop, M.: Seismic inverse scattering in the wave equation approach, MSRI preprint, 2001.

Shen, P., Stolk, C. and Symes, W. W.: Differential semblance velocity analysis by wave-equation migration, Expanded Abstracts, Society of Exploration Geophysicists 73rd Annual International Meeting and Exposition, pp. 2135-2139, 2003.

Mathematical Background

Lions, J. L.: *Nonhomogeneous Boundary Value Problems and Applications*, Springer Verlag, Berlin, 1972.

Nirenberg, L.: *Lectures on Partial Differential Equations*, CMBS Lecture Notes, AMS, 1972.

Treves, F.: *Introduction To Pseudodifferential Operators And Fourier Integral Operators*, Plenum, 1980.

Duistermaat, J.: *Fourier Integral Operators*, Birkhauser, NY, 1996.

Taylor, M.: *Pseudodifferential Operators*, Princeton University Press, 1981.

Hörmander, L.: *The Analysis of Linear Partial Differential Operators*, vols. I-IV, Springer Verlag, 1983.

Geophysical Classics

Hagedoorn, F.: A process of seismic reflection interpretation, *Geophysical Prospecting*, 1954, pp. 85-127.

Claerbout, J.: *Basic Earth Imaging* (“BEI”) and *Imaging the Earth’s Interior* (“IEI”), available through SEP web site, `sepwww.stanford.edu`.

Grau, G. and Versteeg, R.: *The Marmousi Experience: Practical Aspects of Inversion*, IFP-Technip, Paris, 1991.

Dobrin, M. and Savit, C.: *Introduction to Geophysical Prospecting*, 4th edition, McGraw-Hill, 1988.

Yilmaz, O.: *Seismic Data Processing*, 2nd edtion, SEG, 1999.