

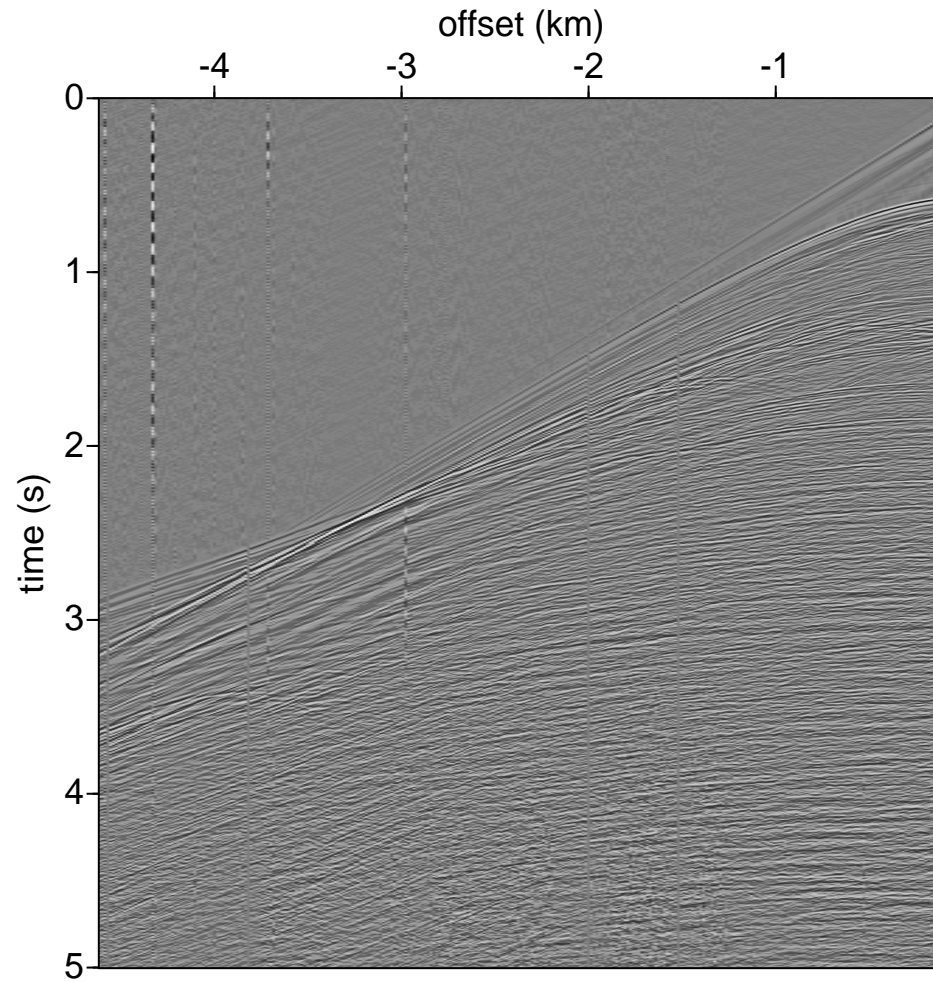
Mathematics of Seismic Imaging

William W. Symes

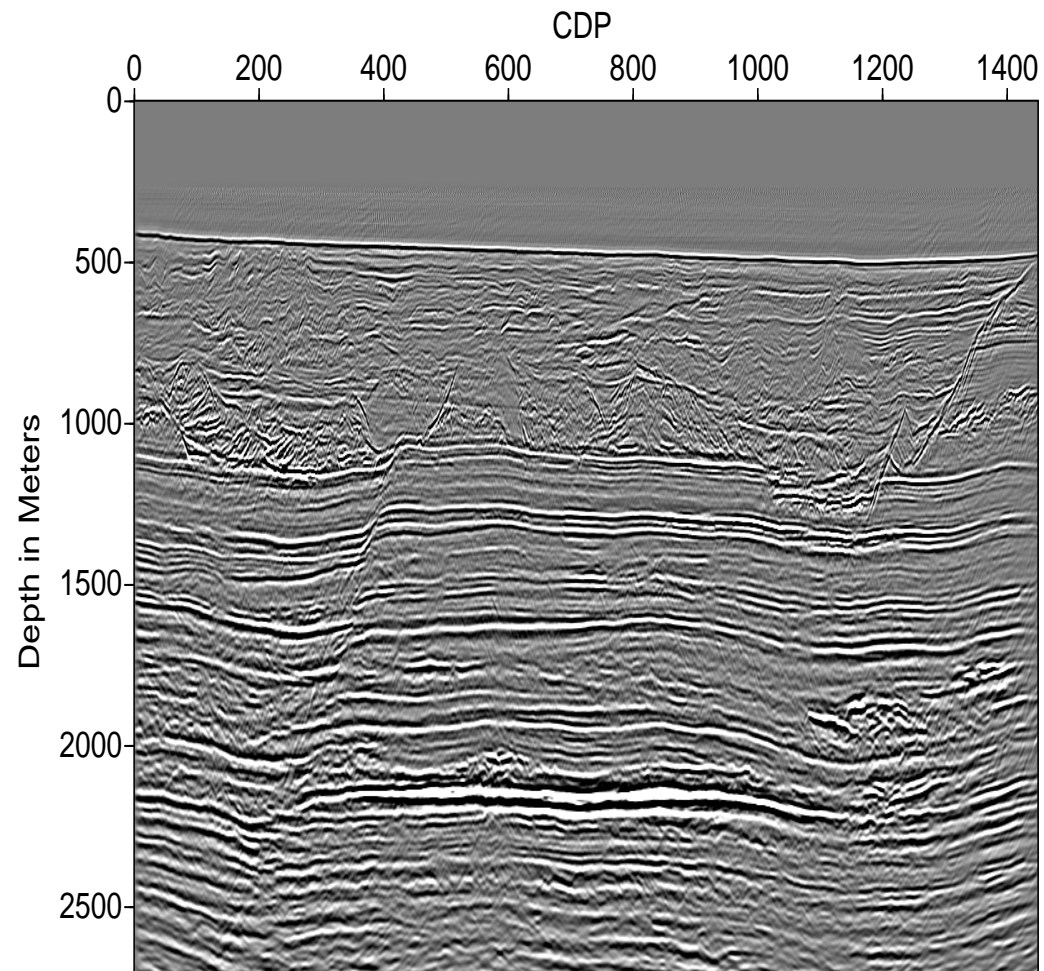
University of Utah

July 2003

www.trip.caam.rice.edu



How do you turn lots of this... (field seismogram from the Gulf of Mexico - thanks: Exxon.)



into this (a fair rendition of subsurface structure)?

Central Point of These Talks:

Estimating the index of refraction (wave velocity) is *the central issue* in seismic imaging.

Combines elements of

- optics, radar, sonar - reflected wave imaging
- tomography - with curved rays

Many unanswered mathematical questions with practical implications!

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

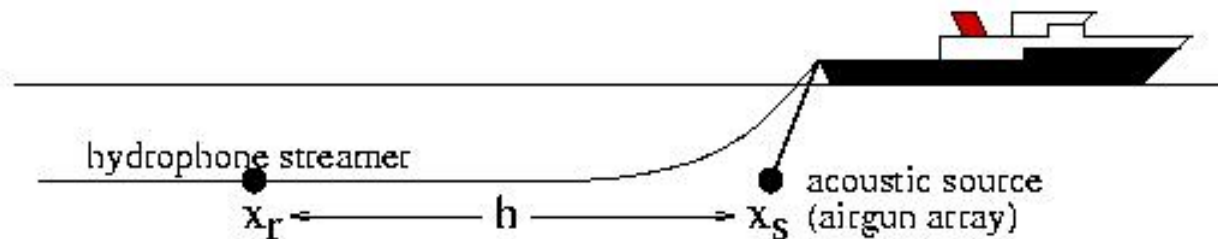
Agenda

1. Seismic inverse problem in the acoustic model: nature of data and model, linearization, reflectors and reflections idealized via *harmonic analysis of singularities*.
2. High frequency asymptotics: why adjoints of modeling operators are imaging operators (“Kirchhoff migration”). Beylkin theory of high frequency asymptotic inversion.
3. Adjoint state imaging with the wave equation: reverse time and reverse depth.
4. Geometric optics, Rakesh’s construction, and asymptotic inversion w/ caustics and multipathing.
5. A step beyond linearization: a mathematical framework for velocity analysis, imaging artifacts, and prestack migration après Claerbout.

1. The Acoustic Model and Linearization

Marine reflection seismology apparatus:

- 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.

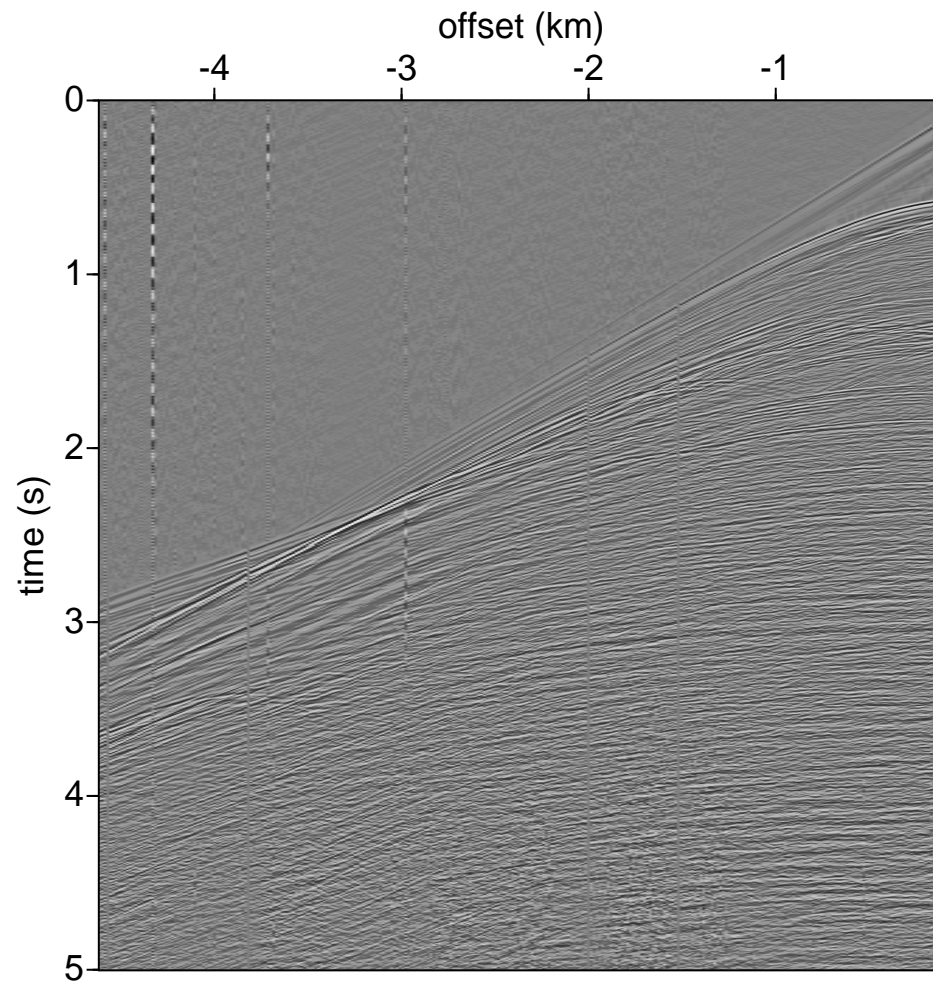
Data parameters: time t , source location x_s , and receiver location x_r or *half offset* $\mathbf{h} = \frac{x_r - x_s}{2}$, $h = |\mathbf{h}|$.

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.]

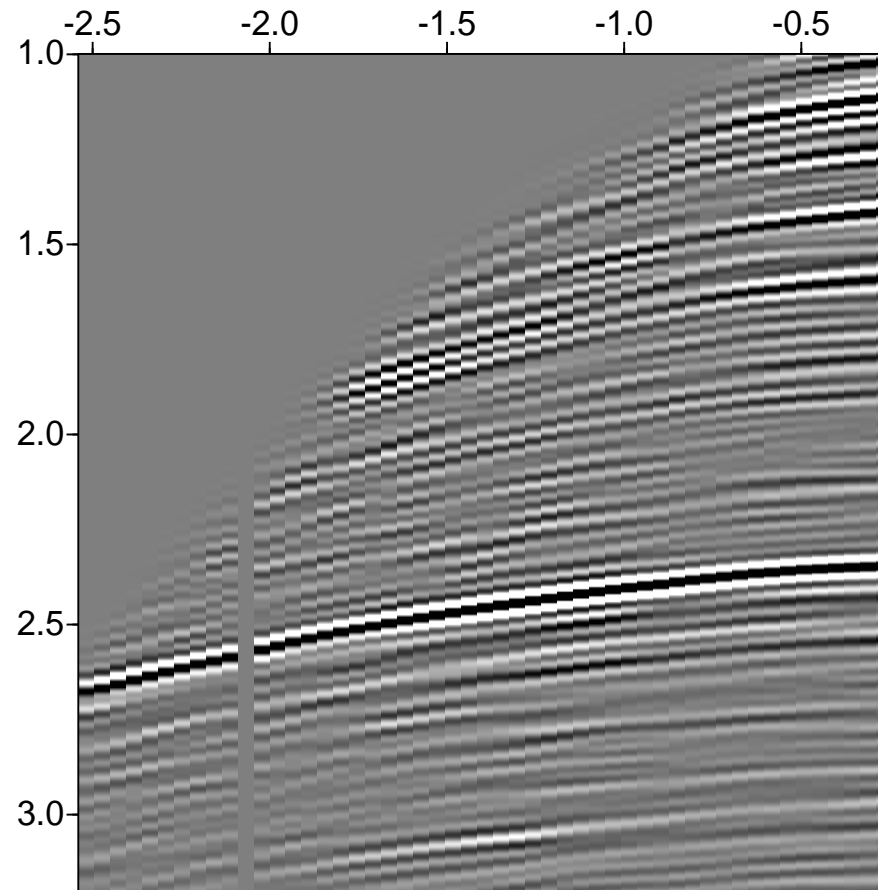
How much data? Contemporary surveys may feature

- 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

Main characteristic of data: *wave nature*, presence of **reflections**
= amplitude coherence along trajectories in space-time.



Data from one source firing, Gulf of Mexico (thanks: Exxon)



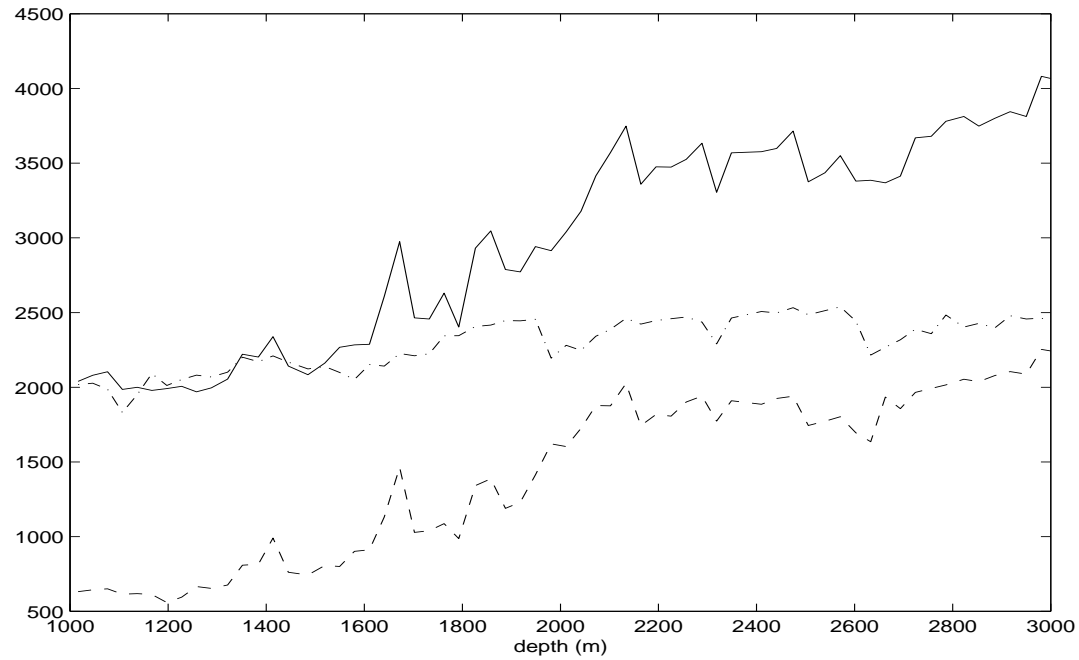
Lightly processed version of data displayed in previous slide - bandpass filtered (in t), truncated (“muted”).

Distinguished data subsets: “gathers”, “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
- ...

- presence of wave events (“reflections”) = coherent space-time structures - clear from examination of the data.
- what features in the subsurface structure could cause reflections to occur?



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/m³). “Blocked” means “averaged” (over 30 m windows). Original sample rate of log tool < 1 m. **Reflectors** = jumps in velocities, density, **velocity trends**.

Known that abrupt (wavelength scale) changes in material mechanics, i.e. reflectors, act as internal boundary, causing reflection of waves.

What is the mechanism through which this occurs?

Seek a simple model which quantitatively explains wave reflection and other known features of the Earth's interior.

The Modeling Task: any model of the reflection seismogram must

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

A *really good* model will also accommodate

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

Acoustic Model (only compressional waves)

Not *really good*, but good enough for today 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):

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

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

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

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

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

Equivalent form: second order wave equation for potential

$$\frac{1}{\rho c^2} \frac{\partial^2 u}{\partial t^2} - \nabla \cdot \frac{1}{\rho} \nabla u = \int_{-\infty}^t dt \nabla \cdot \begin{pmatrix} \mathbf{f} \\ - \\ \rho \end{pmatrix} \equiv \frac{f}{\rho}$$

plus initial, boundary conditions.

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

$$u \in C^1([0, T]; L^2(\Omega)) \cap C^0([0, T]; H_0^1(\Omega))$$

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; initial data

$$u(\cdot, 0) \in H_0^1(\Omega), \quad \frac{\partial u}{\partial t}(\cdot, 0) \in L^2(\Omega)$$

uniquely determine them.

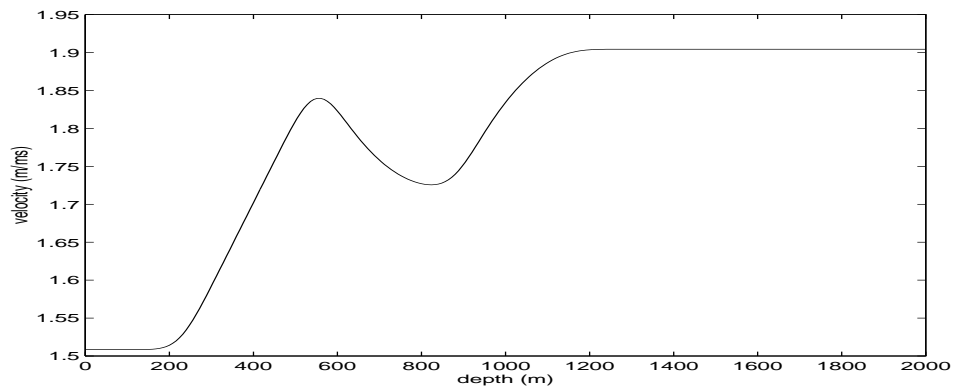
Further idealizations: (i) density is constant, (ii) source force density is *isotropic point radiator with known time dependence* (“source pulse” $w(t)$)

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

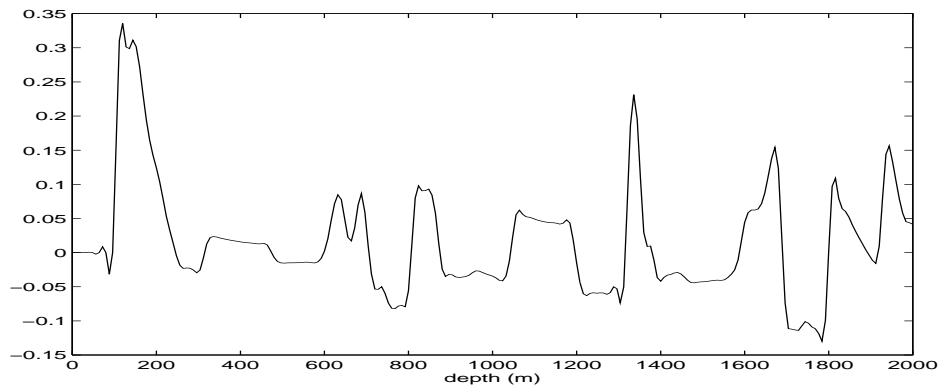
⇒ acoustic potential, pressure depends on \mathbf{x}_s also.

Example: numerical solution of $2D$ wave equation via finite difference method (2nd order in time, 4th order in space).

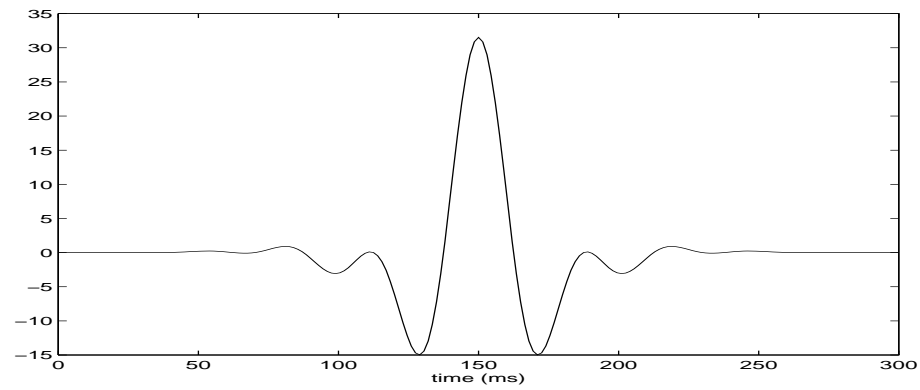
Model = velocity as function of position (in \mathbf{R}^2) = sum of smooth trend plus zero mean fluctuations, *both functions of depth only*. Source wavelet = bandpass filter.



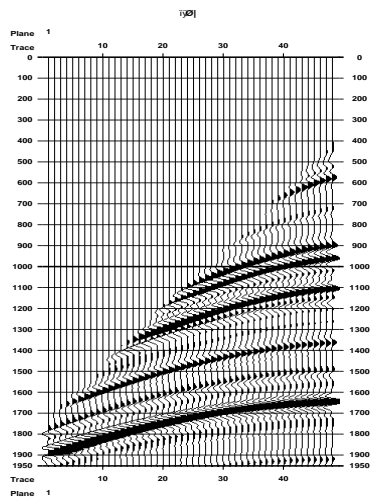
Smooth trend of velocity.



Zero-mean fluctuations of velocity.



4-10-25-35 Hz bandpass filter, used as source wavelet.



Shot (common source) gather or bin, computed by finite difference simulation.

Forward map S = time history of pressure for each \mathbf{x}_s at receiver locations \mathbf{x}_r (predicted seismic data), depends on velocity field $c(\mathbf{x})$:

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

Reflection seismic inverse problem: given *observed seismic data* S^{obs} , find c so that

$$S[c] \simeq S^{\text{obs}}$$

This inverse problem is

- large scale - up to Tbytes, Pflops
- nonlinear

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 pressure 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

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

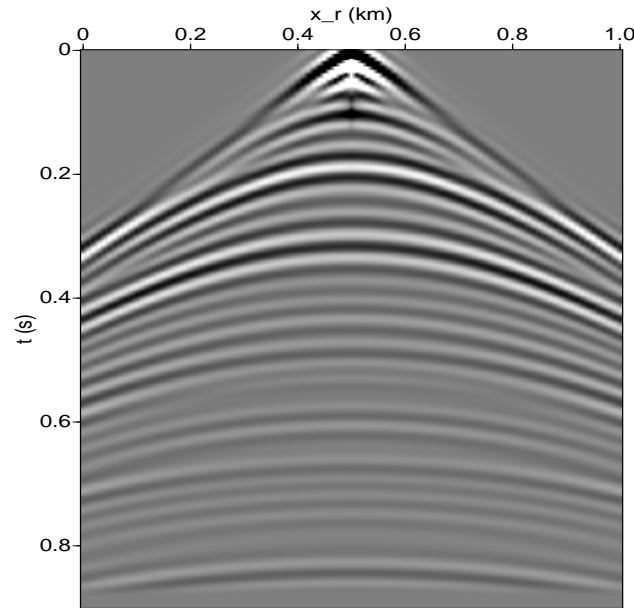
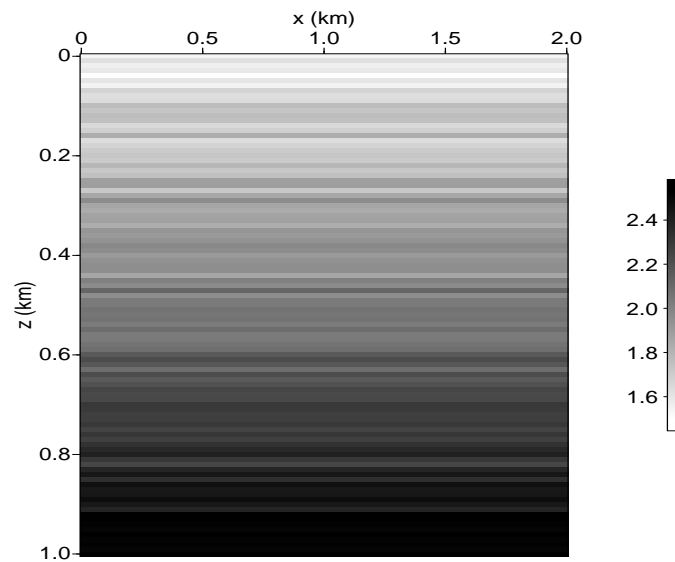
Analysis of $F[v]$ is the main content of contemporary reflection seismic theory.

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

$$S[v(1 + r)] - (S[v] + F[v]r)$$

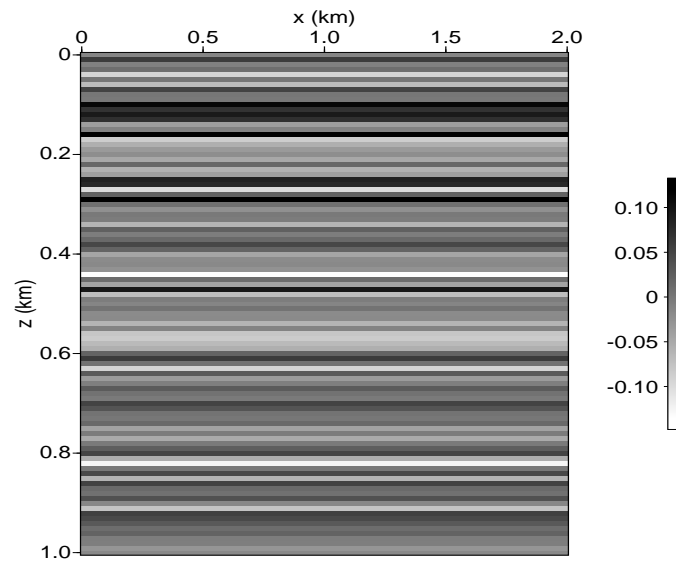
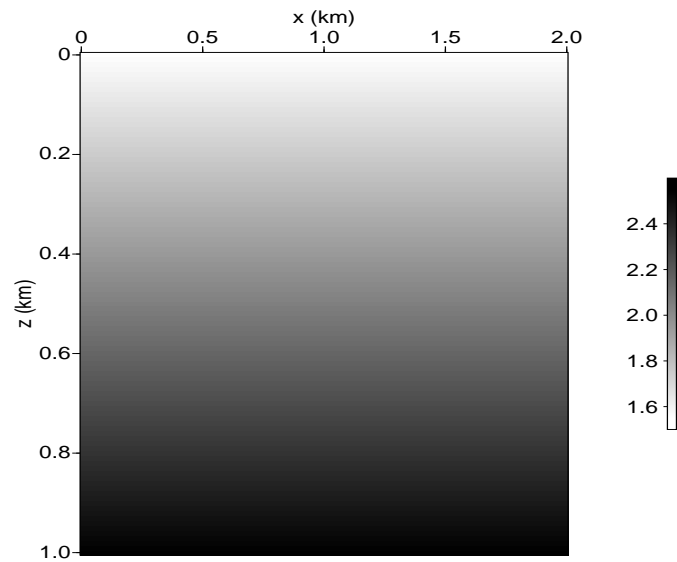
- *small* when v smooth, r rough or oscillatory on wavelength scale - well-separated scales
- *large* when v not smooth and/or r not oscillatory - poorly separated scales

2D finite difference simulation: shot gathers with typical marine seismic geometry. Smooth (linear) $v(x, z)$, oscillatory (random) $r(xz)$ depending only on z (“layered medium”)

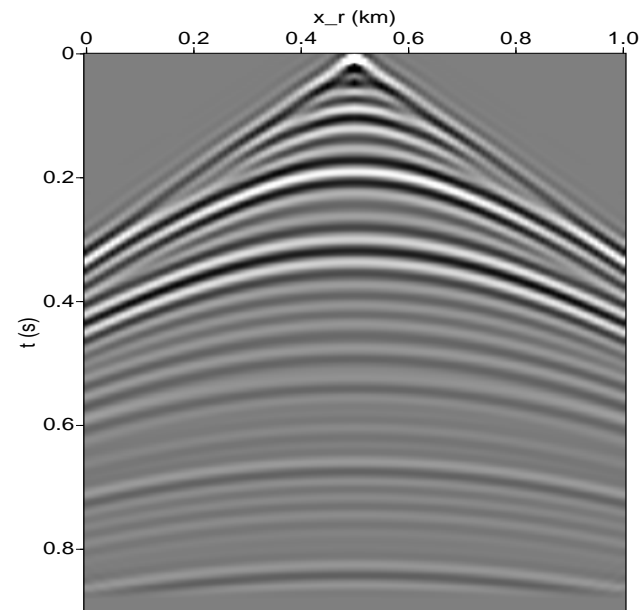
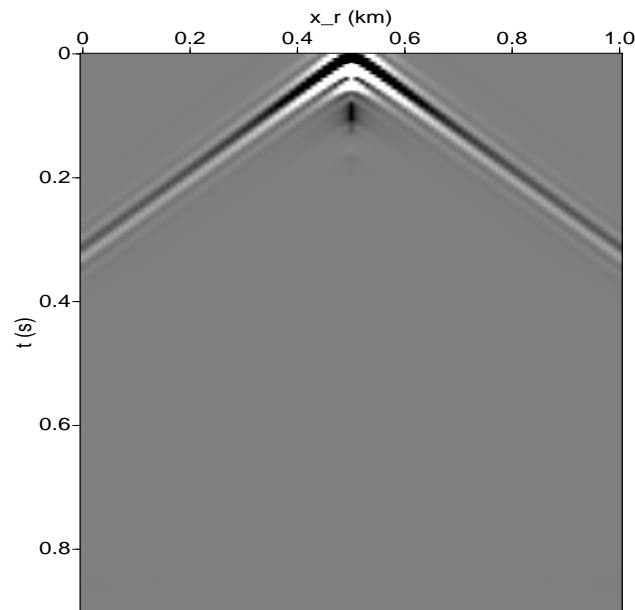


Left: Smooth (linear) background $v(x, z)$, oscillatory (random) $r(x, z)$. Std dev of $r = 5\%$.

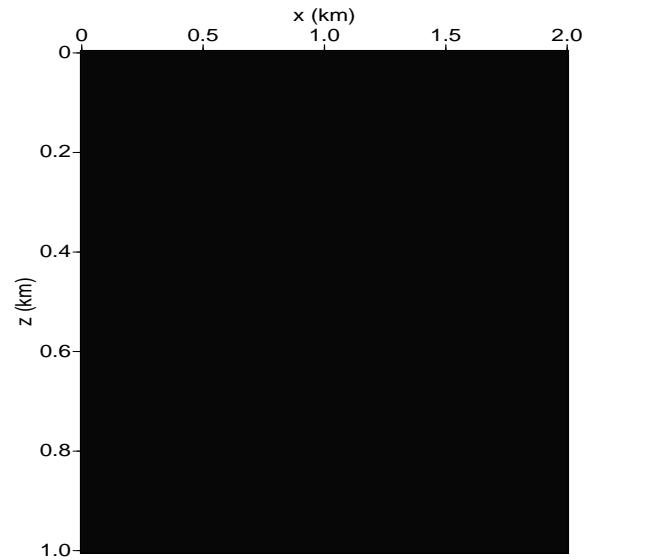
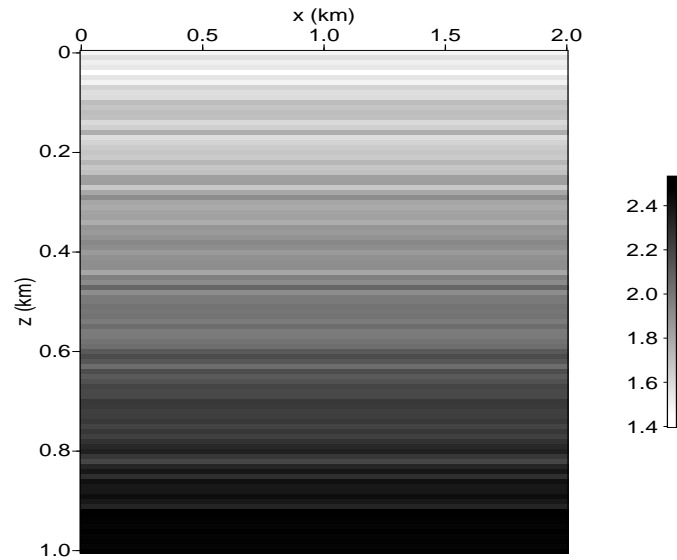
Right: Simulated seismic response ($S[v(1+r)]$), wavelet = band-pass filter 4-10-30-45 Hz. Simulator is (2,4) finite difference scheme.



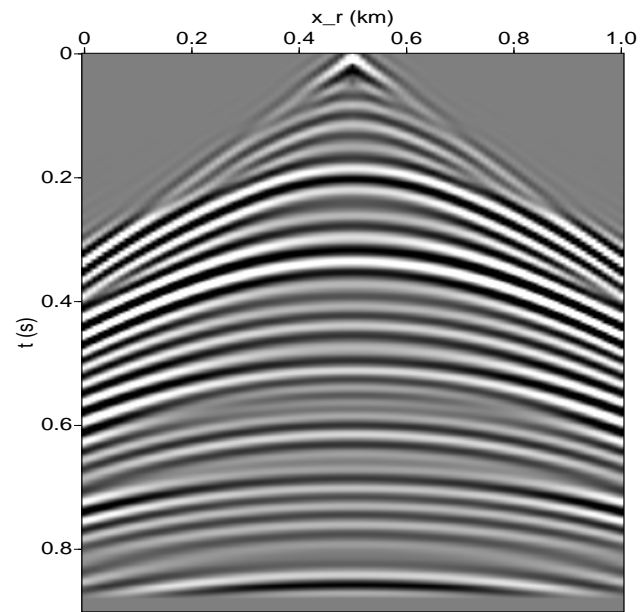
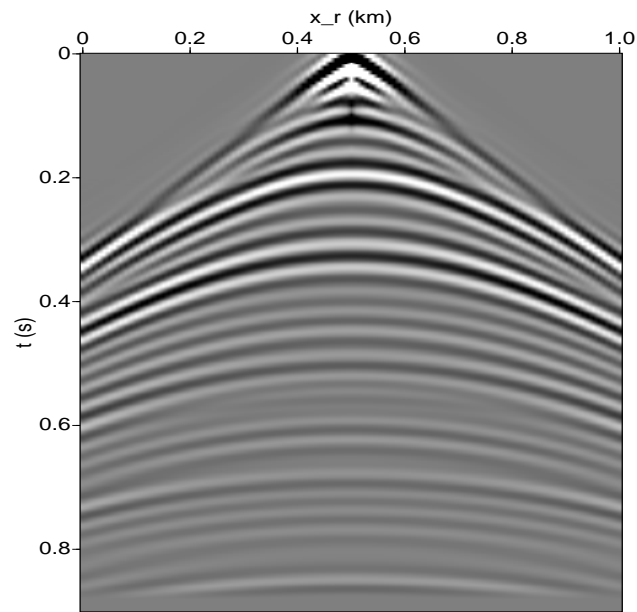
Model in previous slide as smooth background (left, $v(x, z)$) plus rough perturbation (right, $r(x, z)$).



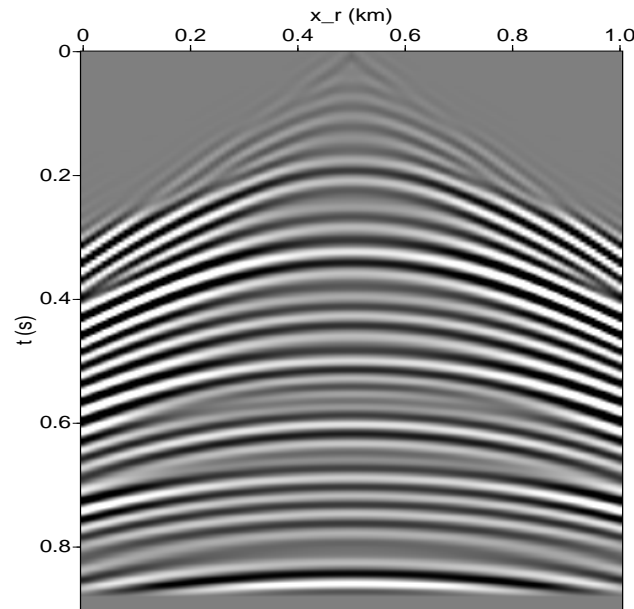
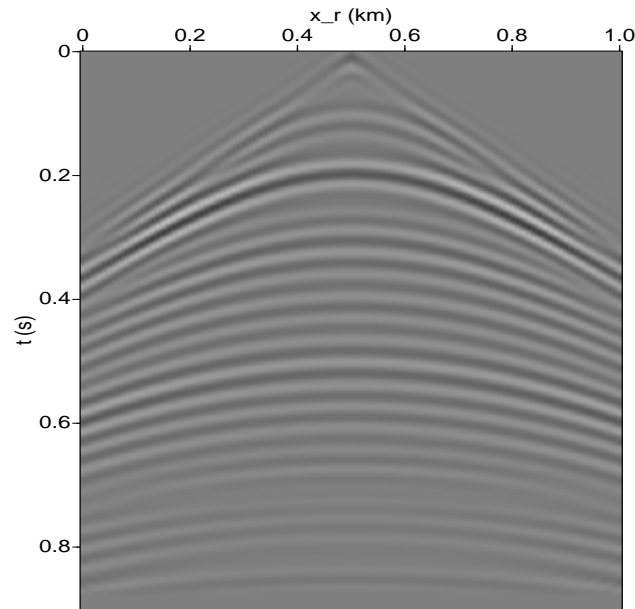
Left: Simulated seismic response of smooth model ($S[v]$),
 Right: Simulated linearized response, rough perturbation of smooth model ($F[v]r$)



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



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



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

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

Implications:

- Some geologies have well-separated scales - cf. sonic logs - linearization-based methods work well there. Other geologies do not - expect trouble!
- 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 (lecture 4).
- 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.

Partially linearized inverse problem = **velocity analysis problem**:
given S^{obs} find v, r so that

$$S[v] + F[v]r \simeq S^{\text{obs}}$$

Linear subproblem = **imaging problem**: given S^{obs} and v , find
 r so that

$$F[v]r \simeq S^{\text{obs}} - S[v]$$

Last 20 years:

- much progress on imaging problem
- much less on velocity analysis problem.

2. High frequency asymptotics and imaging operators

Importance of *high frequency asymptotics*: when linearization is accurate, 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, perhaps in all directions. More flexible notion needed!!

Paley-Wiener characterization of smoothness: $u \in \mathcal{D}'(\mathbf{R}^n)$ is smooth at $\mathbf{x}_0 \Leftrightarrow$ can choose $\phi \in \mathcal{D}(\mathbf{R}^n)$ with $\phi(\mathbf{x}_0) \neq 0$ so that for any N , there is $C_N \geq 0$ so that for any $\xi \neq 0$,

$$|\mathcal{F}(\phi u)(\tau\xi)| \leq C_N(\tau|\xi|)^{-N}, \text{ all } \tau > 0$$

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.

$(\mathbf{x}_0, \xi_0) \notin WF(u) \Leftrightarrow$ can choose $\phi \in \mathcal{D}(\mathbf{R}^n)$ with $\phi(\mathbf{x}_0) \neq 0$ and an open nbhd $\Xi \subset \mathbf{R}^n - 0$ of ξ_0 so that for any N , there is $C_N \geq 0$ so that for all $\xi \in \Xi$,

$$|\mathcal{F}(\phi u)(\tau\xi)| \leq C_N(\tau|\xi|)^{-N}, \text{ all } \tau > 0$$

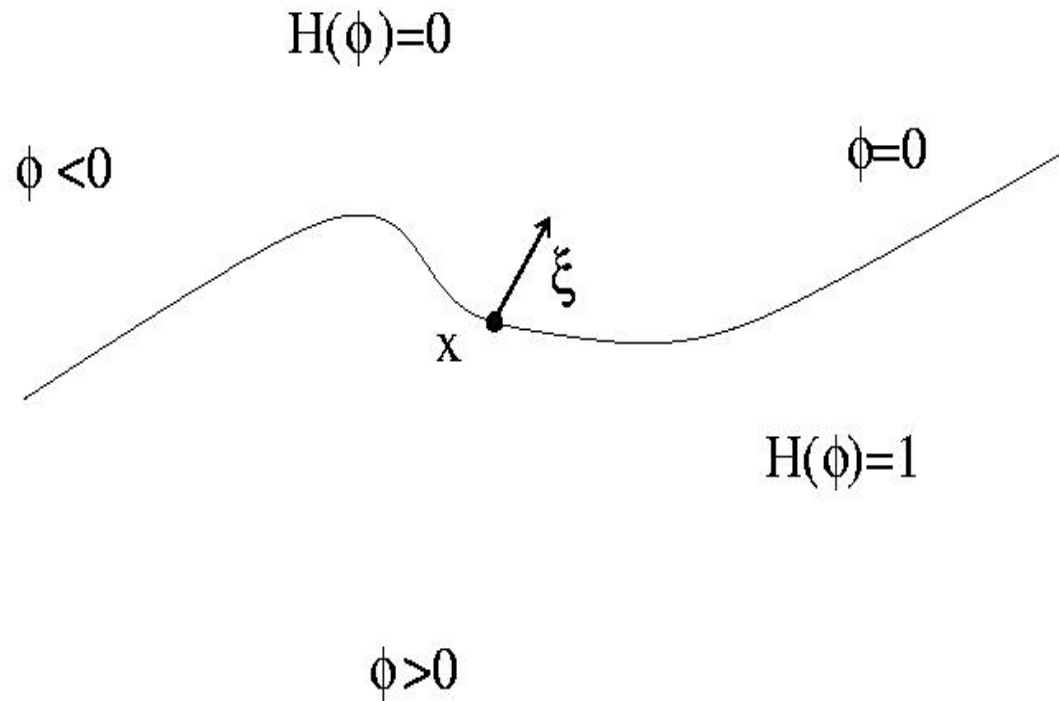
Housekeeping chores:

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

(ii) $WF(u)$ 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 $WF(u) \subset \mathcal{N}_f = \{(\mathbf{x}, \xi) : f(\mathbf{x}) = 0, \xi \parallel \nabla f(\mathbf{x})\}$ (*normal bundle* of $f = 0$).



$$WF(H(f)) = \{(x, \xi) : f(x) = 0, \xi \parallel \nabla f(x)\}$$

Fact (“microlocal property of differential operators”):

Suppose $u \in \mathcal{D}'(\mathbf{R}^n)$, $(\mathbf{x}_0, \xi_0) \notin WF(u)$, and $P(\mathbf{x}, D)$ is a partial differential operator:

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

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

$$\alpha = (\alpha_1, \dots, \alpha_n), \quad |\alpha| = \sum_i \alpha_i,$$

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

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

Proof: Choose $\phi \in \mathcal{D}$, Ξ as in the definition, 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 $\phi_\alpha \in \mathcal{D}$ is a linear combination of derivatives of ϕ and the a_α s. Since each integral is rapidly decreasing as $\tau \rightarrow \infty$ for $\xi \in \Xi$, it remains rapidly decreasing after multiplication by $\tau^{|\alpha|}$, and so does the sum. **Q. E. D.**

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)

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

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

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

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

Geometric optics approximation of G should be good, as v is smooth. Local version: 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 travelttime $\tau(\mathbf{x}; \mathbf{x}_s)$ solves the eikonal equation

$$v|\nabla\tau| = 1$$

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

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

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

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

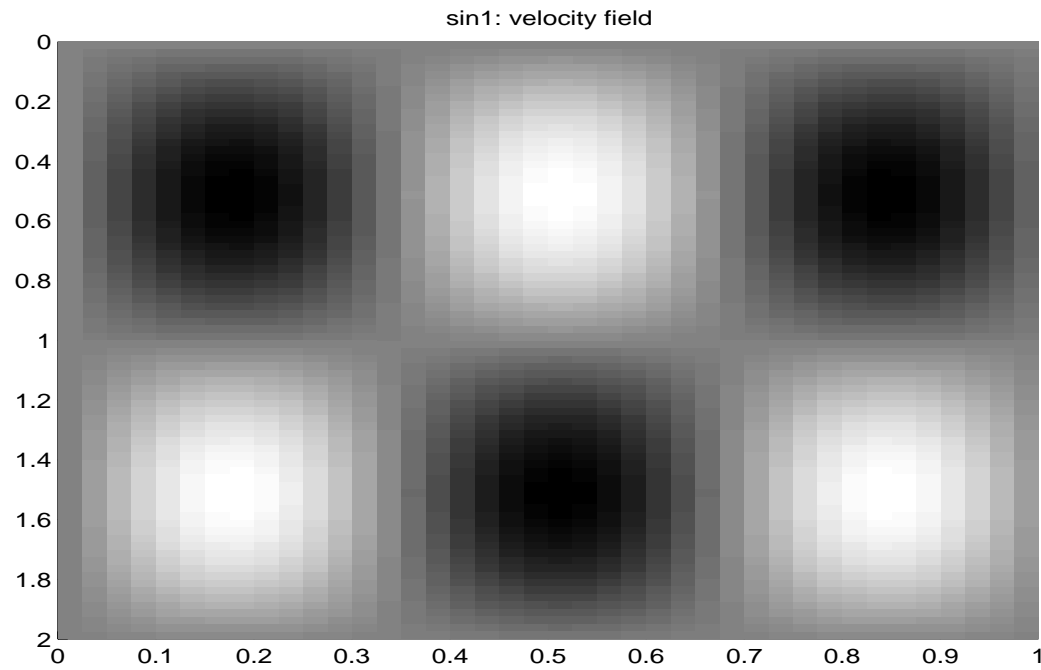
$$\int dx \int_0^T dt |R(\mathbf{x}, t; \mathbf{x}_s)|^2 \leq C \|v\|_{C^4}$$

Numerical solution of eikonal, transport: ray tracing (Lagrangian), various sorts of upwind finite difference (Eulerian) methods. See Sethian lectures, WWS 1999 MGSS notes (online) for details.

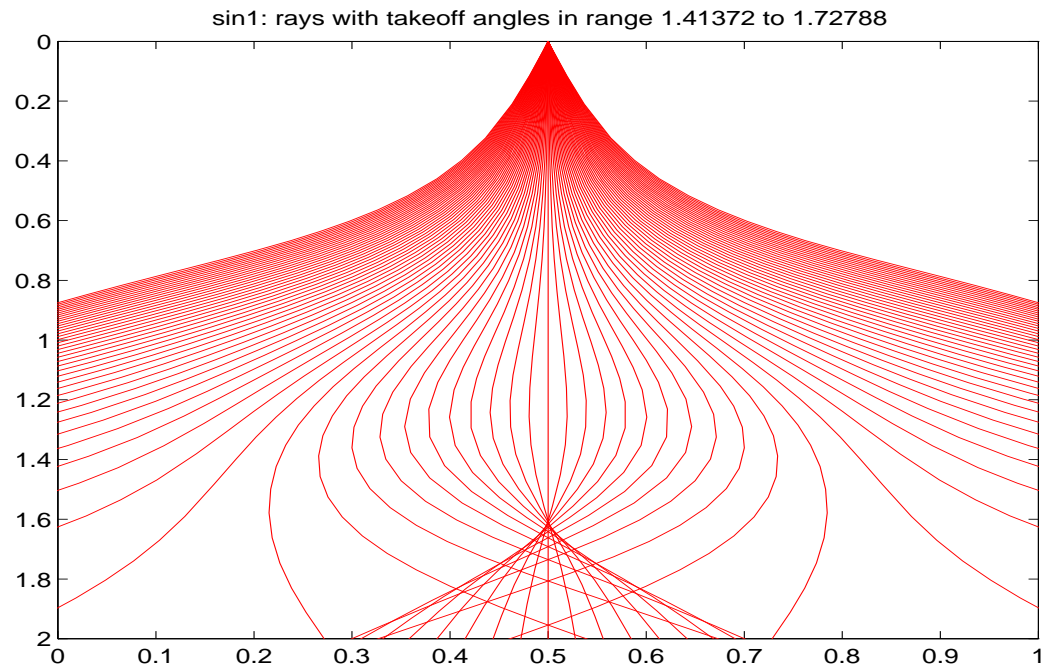
“Not too far” means: there should be one and only one ray of geometric optics connecting each \mathbf{x}_s or \mathbf{x}_r to each $\mathbf{x} \in \text{supp}r$.

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* (White, 1982).

Upon caustic formation, the simple geometric optics field description above is no longer correct (Ludwig, 1966).



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



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

Assume: $\text{supp } r$ contained in simple geometric optics domain (each point reached by unique ray from any source point \mathbf{x}_s).

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

$$\begin{aligned} K(\mathbf{x}_r, t, \mathbf{x}_s; \mathbf{x}) &= \int ds G(\mathbf{x}_r, t - s; \mathbf{x}) \frac{\partial^2 G}{\partial t^2}(\mathbf{x}, s; \mathbf{x}_s) \frac{2}{v^2(\mathbf{x})} \\ &\simeq \int ds \frac{2a(\mathbf{x}_r, \mathbf{x})a(\mathbf{x}, \mathbf{x}_s)}{v^2(\mathbf{x})} \delta'(t - s - \tau(\mathbf{x}_r, \mathbf{x})) \delta''(s - \tau(\mathbf{x}, \mathbf{x}_s)) \\ &= \frac{2a(\mathbf{x}, \mathbf{x}_r)a(\mathbf{x}, \mathbf{x}_s)}{v^2(\mathbf{x})} \delta''(t - \tau(\mathbf{x}, \mathbf{x}_r) - \tau(\mathbf{x}, \mathbf{x}_s)) \end{aligned}$$

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].

Q: What does \simeq mean?

A: It means “differs by something smoother”.

In theory, 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.

In practice, it's sufficient to make the difference just a bit smoother, so the first term of the geometric optics approximation (displayed above) suffices (can formalize this with modification of wavefront set defn).

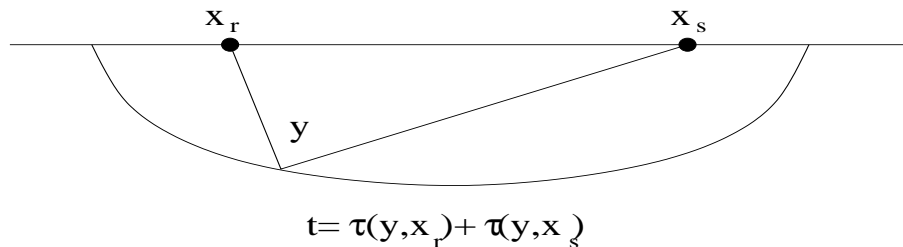
These lectures will ignore the distinction.

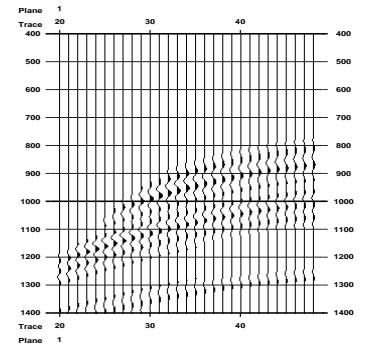
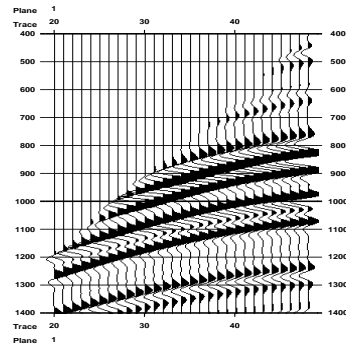
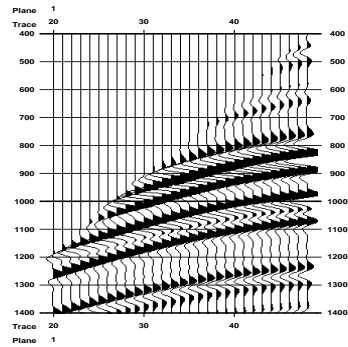
So: for r supported in simple geometric optics domain, no forward scattering \Rightarrow

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

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

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 $v = \text{const.}$ then isochron is ellipsoid, as $\tau(\mathbf{x}_s, \mathbf{x}) = |\mathbf{x}_s - \mathbf{x}|/v!$





Left: FD *linearized* simulation of shot gather, (i.e. $\delta G(\mathbf{x}_r, t; \mathbf{x}_s)$ for fixed \mathbf{x}_s , convolved with bandpass filter source wavelet).

Middle: Detail of numerical implementation of geometric optics approximation.

Right: Difference, plotted on same scale.

Zero Offset data and the Exploding Reflector

Zero offset data ($x_s = x_r$) 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

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

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 travelttime, 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, BEI, for examples which demonstrate that multipathing really does invalidate exploding reflector model.

Standard Processing: an inspirational interlude

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 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") = traces with the same *midpoint* $\mathbf{x}_m = (\mathbf{x}_r + \mathbf{x}_s)/2$.

This isn't really true - *look at the data!!!* However it is *approximately* correct in many places in the world: CMPs change slowly with midpoint

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 $v = v(\mathbf{x}_m, z), r = r(\mathbf{x}_m, z)$.

$$\begin{aligned}
 & F[v]r(\mathbf{x}_r, t; \mathbf{x}_s) \\
 & \simeq \int dx \frac{2r(z)}{v^2(z)} a(\mathbf{x}, x_r) a(\mathbf{x}, x_s) \delta''(t - \tau(\mathbf{x}, x_r) - \tau(\mathbf{x}, x_s)) \\
 & = \int dz \frac{2r(z)}{v^2(z)} \int d\omega \int dx \omega^2 a(\mathbf{x}, x_r) a(\mathbf{x}, x_s) e^{i\omega(t - \tau(\mathbf{x}, x_r) - \tau(\mathbf{x}, x_s))}
 \end{aligned}$$

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,

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

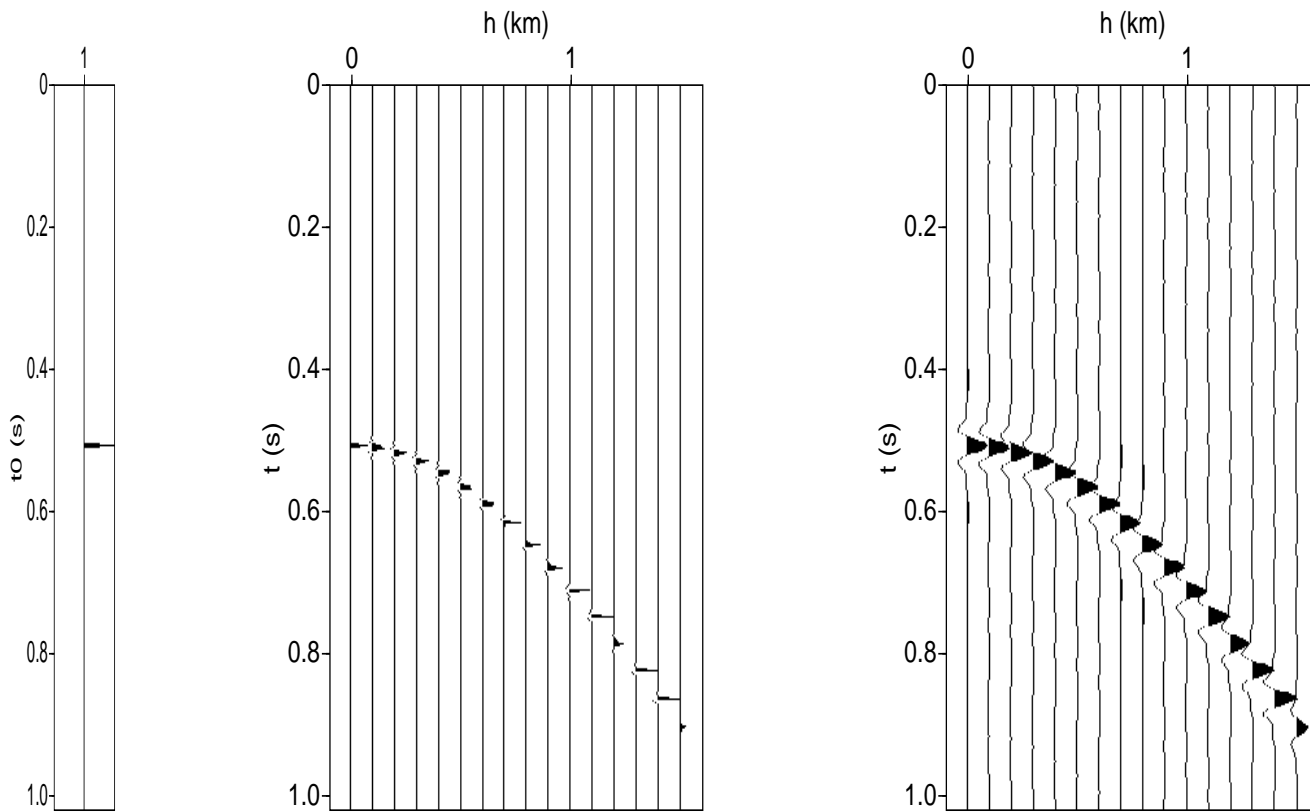
Here $z(h, t)$ is the inverse of the 2-way traveltimes

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

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 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.



Left: $r(t_0)$. Middle: $r(\zeta(t, h))$. Right: $F[v]r(t, h)$

Anatomy of an adjoint:

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

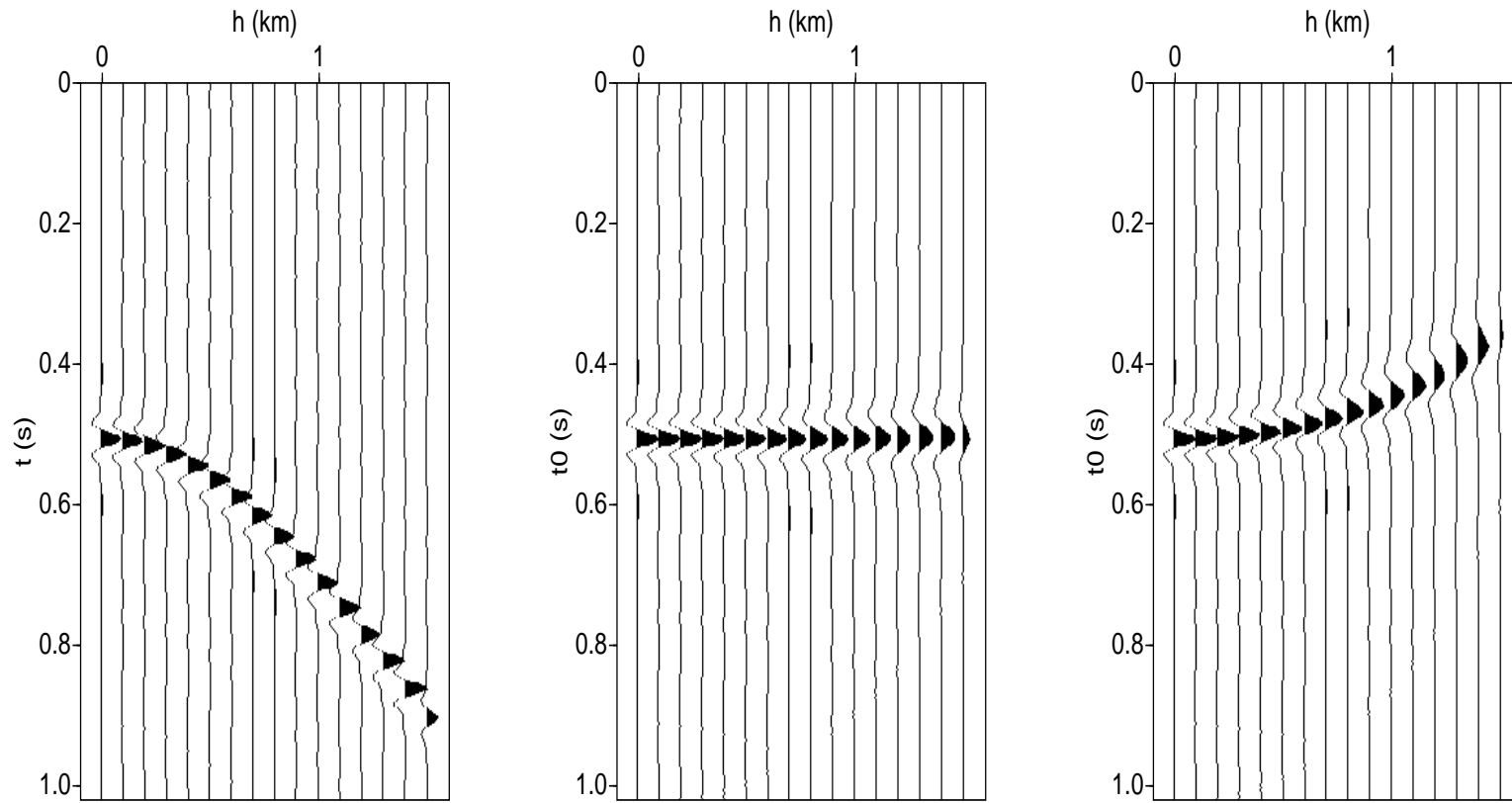
$$\text{so } F[v]^* = -\frac{\partial}{\partial z} S M[v] N[v],$$

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

$$M[v] = \text{multiplication by } \frac{\partial t}{\partial z} A$$

$$S = \mathbf{stacking\ operator}$$

$$S f(z) = \int dh f(z, h)$$



Left: $d(t, h) = F[v']r(t, h)$. Middle: $N[v']d(t_0, h)$. Right:
 $N[v_1]d(t_0, h), v_1 \neq v$

Stack as ersatz ZO section: how to "get" ZO data even if you don't measure it.

If you use t_0 instead of z to parametrize depth, *and* data is consistent with conv. model, then

every trace looks exactly like the zero offset trace - even if you don't measure the zero offset trace!

Therefore stack = average of a bunch of traces (output of NMO) which have events in exactly the same place (t_0) as ZO trace. Hence view stack as substitute for ZO data.

Adjusting velocity estimate to align NMO corrected traces produces *stacking velocity* - if successful, then stack enhances signal, suppresses (coherent and incoherent) noise.

Imaging in standard processing: suppose that S^{obs} is model-consistent, i.e. $d \equiv S^{\text{obs}} - S[v] = F[v]r$. Then

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

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!

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

Multioffset Imaging: if $d = F[v]r$, then

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

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.*

Beylkin, JMP 1985: for r supported in simple geometric optics domain,

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

Also: explicit relation between $WF(d)$ and $WF(r)$ ("canonical relation").

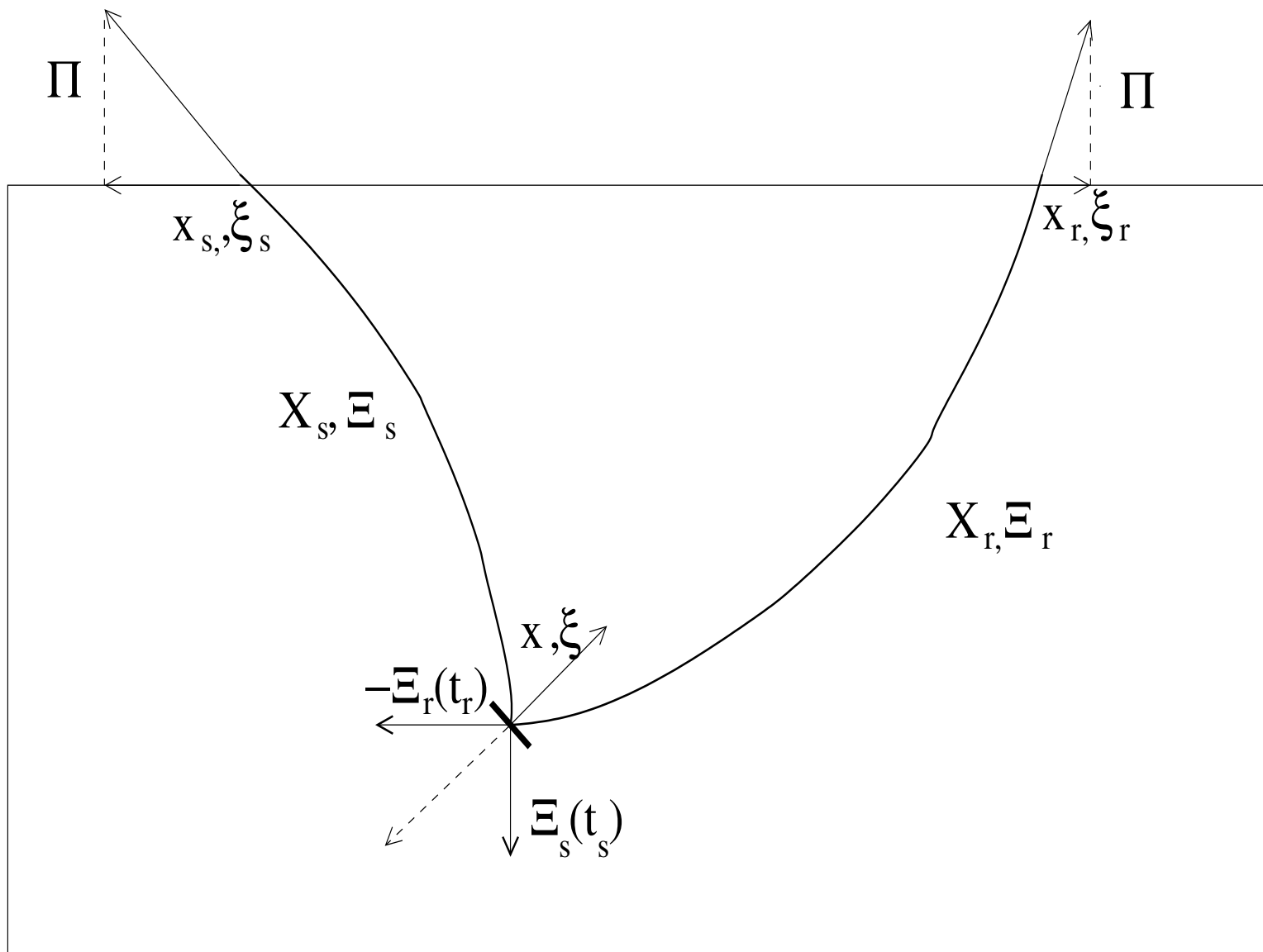
$$(\mathbf{x}, \xi) \in WF(r) \Leftrightarrow (\mathbf{x}_s, \mathbf{x}_r, t, \xi_s, \xi_r, \omega) \in WF(d)$$

and (1) the source ray connects (\mathbf{x}_s, ξ_s) to (\mathbf{x}, Ξ_s) with traveltime $\tau(\mathbf{x}, \mathbf{x}_s)$;

(2) the receiver ray connects (\mathbf{x}_r, ξ_r) to (\mathbf{x}, Ξ_r) with traveltime $\tau(\mathbf{x}, \mathbf{x}_r)$;

(3) $\xi \parallel \Xi_s + \Xi_r$ ("Snell's law");

(4) $t = \tau(\mathbf{x}, \mathbf{x}_s) + \tau(\mathbf{x}, \mathbf{x}_r)$.



Outline of proof: (i) 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* (“ Ψ DO”):

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

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

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

Explicit computation of **symbol** p - for details, .

Imaging property of Kirchhoff migration follows from *microlocal property of Ψ DOs*:

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

Will prove this. First, a few other properties:

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

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

Proof of microlocal property: suppose $(\mathbf{x}_0, \xi_0) \notin WF(u)$, 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 $B_{3\epsilon}(\xi_0/|\xi_0|) \subset \Xi$, and chose $\Xi' = \{\tau\xi : \xi \in B_\epsilon(\xi_0/|\xi_0|), \tau > 0\}$.

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.

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

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

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

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

using the identity

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

and integrating by parts $2M$ times in ξ . 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 ,

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

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:

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

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

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

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

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

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

$$\begin{aligned} \tau^{n-2M} \left| \int dx \int_{|\theta - \eta| > \epsilon} d\theta (-\nabla_x^2)^M [P(\mathbf{x}, \tau\theta)\phi(\mathbf{x})] \widehat{\psi} u(\tau\theta) \right. \\ \left. \times |\theta - \eta|^{-2M} e^{i\tau \mathbf{x} \cdot (\theta - \eta)} \right| \\ \leq C \tau^{n+m-2M} \end{aligned}$$

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.

Recall: in layered case,

$$F[v]r(h, t) \simeq A(z(h, t), h) \frac{1}{2} \frac{dr}{dz}(z(h, t))$$

$$F[v]^* d(z) \simeq -\frac{\partial}{\partial z} \int dh A(z, h) \frac{\partial t}{\partial z}(z, h) d(t(z, h), h)$$

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

thus normal operator is invertible and you can construct approximate least-squares solution to $F[v]r = d$:

$$\tilde{r} \simeq (F[v]^* F[v])^{-1} F[v]^* d$$

Relation between r and \tilde{r} : difference is *smoother* than either. Thus difference is *small* if r is oscillatory - consistent with conditions under which linearization is accurate.

Analogous construction in simple geometric optics case: due to Beylkin (1985).

Complication: $F[v]^*F[v]$ cannot be invertible - because $WF(F[v]^*F[v]r)$ generally quite a bit smaller than $WF(r)$.

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

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

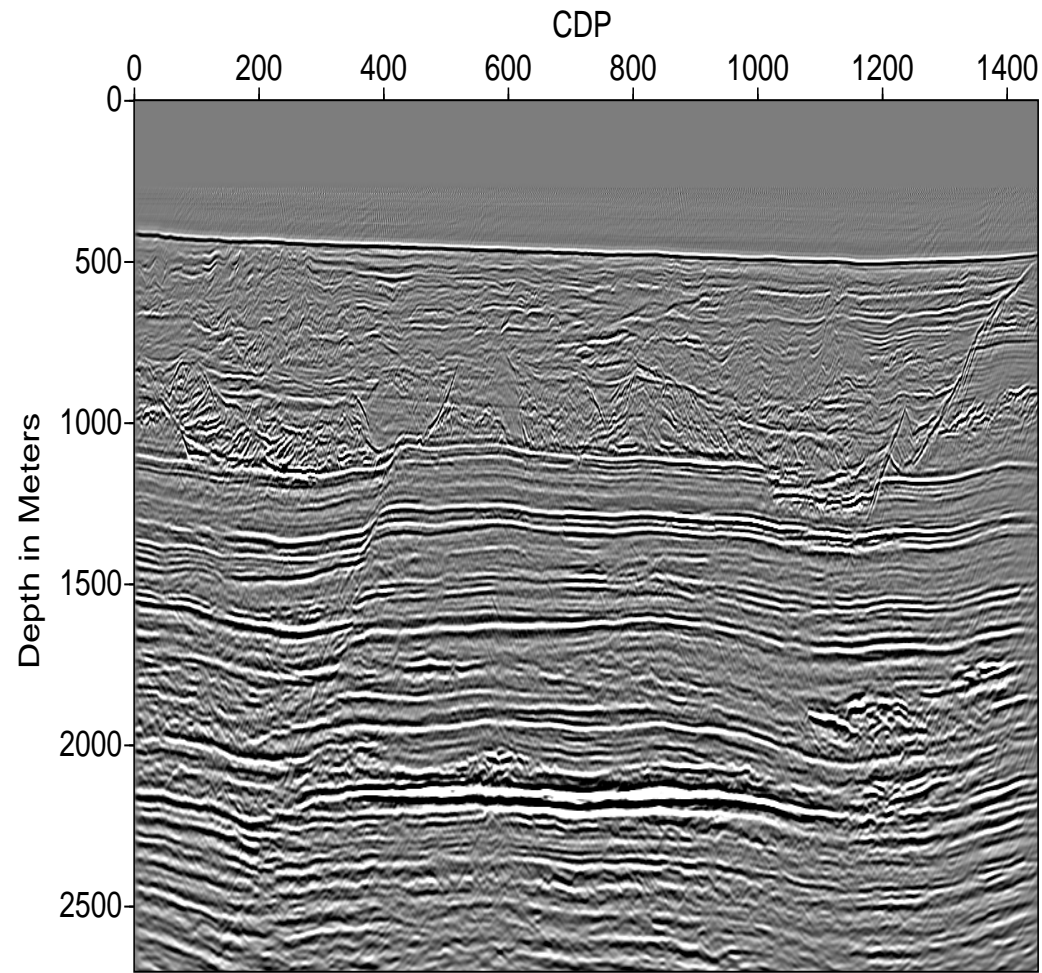
yields $F[v]^\dagger F[v]r \simeq r$ if $WF(r) \subset \Gamma[v]$

For details of Beylkin construction: Beylkin, 1985; Miller et al 1989; Bleistein, Cohen, and Stockwell 2000; WWS 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) - how much of this can be rescued? cf. Lecture 3.



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.