Mathematics of Seismic Imaging Part 2: Linearization, High Frequency Asymptotics, and Imaging

William W. Symes

Rice University

2. Linearization, High frequency Asymptotics and Imaging

2. Linearization, High frequency Asymptotics and Imaging

Linearization

All useful technology relies somehow on linearization (aka perturbation theory, Born approximation,...): write $c = v(1+r)$, $r =$ relative first order perturbation about $v \Rightarrow$ perturbation of pressure field $\delta p = \frac{\partial \delta u}{\partial t} = 0, t \leq 0$,

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

linearized forward map F:

$$F[v]r = \delta p|_{\Sigma\times[0,T]}$$

Linearization in theory

Recall Lions-Stolk result: if $\log c \in L^\infty(\Omega)$ ($\rho = 1$!) and $f \in L^2(\Omega \times [0, T])$, then weak solution has *finite energy*, i.e.

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

Suppose $\delta c \in L^\infty(\Omega)$, define δu by solving perturbational problem: set $v = c$, $r = \delta c / c$.

Linearization in theory

Stolk (2000): for $\delta c \in L^\infty(\Omega)$, small enough $h \in \mathbf{R}$,

$$\|u[c + h\delta c] - u[c] - \delta u\|_{C^0([0,T],L^2(\Omega))} = o(h)$$

Note "loss of derivative": error in Newton quotient is $o(1)$ in weaker norm than that of space of weak solns

Linearization in theory

Implication for $\mathcal{F}[c]$: under suitable circumstances ($c =$ const. near Σ - "marine" case),

$$\|\mathcal{F}[c]\|_{L^2(\Sigma\times[0,T])} = O(\|w\|_{L^2(\mathbf{R})})$$

but

$$\|\mathcal{F}[v(1+r)]-\mathcal{F}[v]-F[v]r\|_{L^2(\Sigma\times[0,T])} = O(\|w\|_{H^1(\mathbf{R})})$$

and these estimates are both sharp

Linearization in practice

Physical intuition, numerical simulation, and not nearly enough mathematics: linearization error

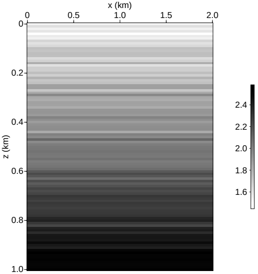
$$\mathcal{F}[v(1+r)] - \mathcal{F}[v] - F[v]r$$

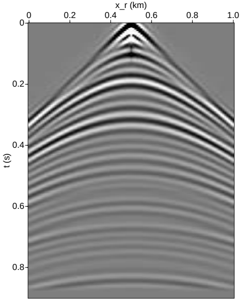
- *small* when v smooth, r rough or oscillatory on wavelength scale - well-separated scales
- *large* when v not smooth and/or r not oscillatory - poorly separated scales

RICE

Linearization in practice

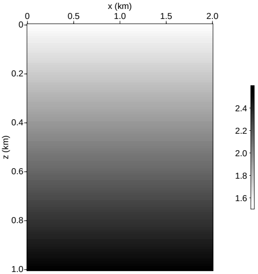
Illustration: 2D finite difference simulation: shot gathers with typical marine seismic geometry. Smooth (linear) $v(x,z)$, oscillatory (random) $r(x,z)$ depending only on z("layered medium"). Source wavelet $w(t) =$ bandpass filter.



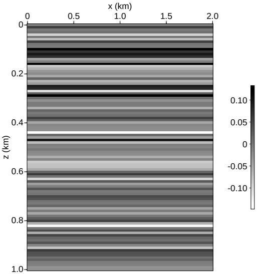


,

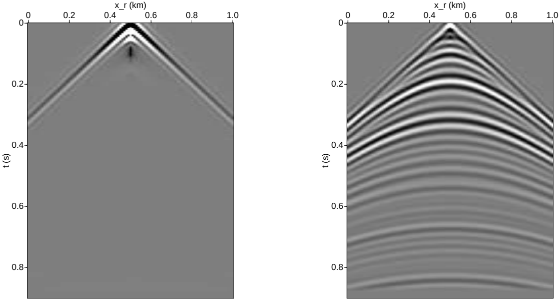
Left: $c = v(1+r)$. Std dev of $r = 5\%$.
Right: Simulated seismic response $(\mathcal{F}[v(1+r)])$, wavelet = bandpass filter 4-10-30-45 Hz. Simulator is (2,4) finite difference scheme.



,

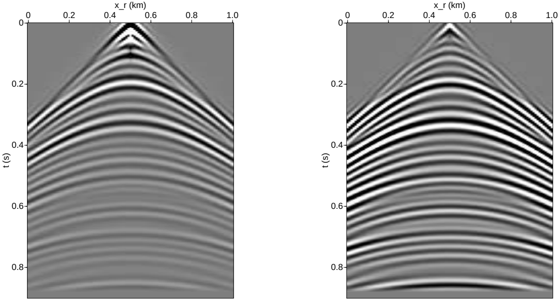


Decomposition of model in previous slide as smooth background (left, $v(x,z)$) plus rough perturbation (right, $r(x,z)$).



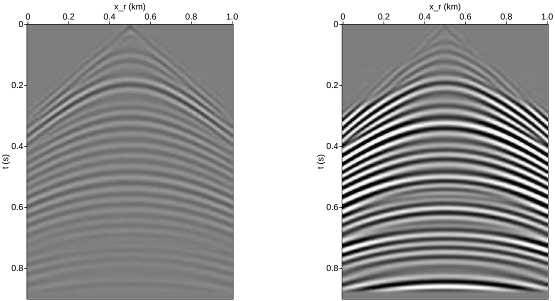
.

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



.

Left: Simulated seismic response of rough model ($\mathcal{F}[0.95v + r]$),
Right: Simulated linearized response, smooth perturbation of rough model ($F[0.95v + r]((0.05v)/(0.95v + r))$)



,

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

For the same pulse w,

- v smooth, r oscillatory $\Rightarrow F[v]r$ approximates **primary reflection** = result of one-time wave-material interaction (single scattering); error = **multiple reflections**, "not too large" if r is "not too big"
- v nonsmooth, r smooth $\Rightarrow$ error = *time shifts* - very large perturbations since waves are oscillatory.

For typical oscillatory w ($\|w\|_{H^1} >> \|w\|_{L^2}$), tends to imply that *in scale-separated case, effectively no loss of derivative!*

2. Linearization, High frequency Asymptotics and Imaging

Velocity Analysis and Imaging

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

$$\mathcal{F}[v] + F[v]r \simeq d$$

Linearized inversion problem: given d and v, find r so that

$$F[v]r \simeq d - \mathcal{F}[v]$$

Imaging problem - relaxation of linearized inversion: given d and v, find an *image* r of "reality" = solution of linearized inversion problem

Velocity Analysis and Imaging

Last 20 years: mathematically speaking,

- much progress on imaging
- lots of progress on linearized inversion
- much less on velocity analysis
- none to speak of on nonlinear inversion

[Caveat: a *lot* of practical progress on nonlinear inversion in the last 10 years!]

Velocity Analysis and Imaging

Interesting question: what's an image?

"...I know it when I see it." - Associate Justice Potter Stewart, 1964

2. Linearization, High frequency Asymptotics and Imaging

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.

Aymptotic assumption

NB: "everybody's favorite reflector": the smooth interface across which r jumps.

But this is an oversimplification - waves reflect at complex zones of rapid change in rock mechanics, pehaps in all directions. More flexible notion needed!!

RICE

Wave Front Set

Paley-Wiener characterization of local smoothness for distributions: $u \in \mathcal{D}'(\mathbf{R}^n)$ is smooth at $\mathbf{x}_0 \Leftrightarrow$ for some nbhd X of $\mathbf{x}_0$, any $\chi \in C_0^\infty(X)$ and $N \in \mathbf{N}$, any $\boldsymbol{\xi} \in \mathbf{R}^n, |\boldsymbol{\xi}| = 1$,

$$|\widehat{(\chi u)}(\tau\boldsymbol{\xi})| = O(\tau^{-N}),\ \tau \rightarrow \infty$$

Proof (sketch): smooth at $\mathbf{x}_0$ means: for some nbhd X, $\chi u \in C_0^\infty(\mathbf{R}^n)$ for any $\chi \in C_0^\infty(X) \Leftrightarrow$.

$$\widehat{\chi u}(\boldsymbol{\xi}) = \int dx\, e^{i\boldsymbol{\xi}\cdot\mathbf{x}}\chi(x)u(x)$$

Wave Front Set

$$= \int dx\,(1+|\boldsymbol{\xi}|^2)^{-p}[(I-\nabla^2)^p e^{i\boldsymbol{\xi}\cdot\mathbf{x}}]\chi(x)u(x)$$

$$= (1+|\boldsymbol{\xi}|^2)^{-p}\int dx\, e^{i\boldsymbol{\xi}\cdot\mathbf{x}}[(I-\nabla^2)^p\chi(x)u(x)]$$

whence

$$|\widehat{\chi u}(\boldsymbol{\xi})| \le \mathrm{const.}(1+|\boldsymbol{\xi}|^2)^{-p}$$

where the const. depends on p, χ and u. For any N, choose p large enough, replace $\boldsymbol{\xi} \leftarrow \tau\boldsymbol{\xi}$, get desired $\le$.

Wave Front Set

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

$(\mathbf{x}_0, \boldsymbol{\xi}_0) \notin WF(u) \Leftrightarrow$, there is open nbhd $X \times \Xi \subset \mathbf{R}^n \times \mathbf{R}^n \setminus 0$ of $(\mathbf{x}_0, \boldsymbol{\xi}_0)$ so that for any $\chi \in C_0^\infty(\mathbf{R}^n), \mathrm{supp}\chi \subset X$, $N \in \mathbf{N}$, all $\boldsymbol{\xi} \in \Xi$ so that $|\boldsymbol{\xi}| = |\boldsymbol{\xi}_0|$,

$$|\widehat{\chi u}(\tau \boldsymbol{\xi})| = O(\tau^{-N})$$

Housekeeping chores

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

(ii) $WF(u)$ is invariant under chg. of coords - as subset of the *cotangent bundle* $T^*(\mathbf{R}^n)$ (i.e. the $\boldsymbol{\xi}$ components transform as covectors).

(iii) Standard example: if u jumps across the interface $\phi(\mathbf{x}) = 0$, otherwise smooth, then $WF(u) \subset \mathcal{N}_\phi = \{(\mathbf{x}, \boldsymbol{\xi}) : \phi(\mathbf{x}) = 0, \, \boldsymbol{\xi} || \nabla\phi(\mathbf{x})\}$ (*normal bundle* of $\phi = 0$)

[Good refs for basics on WF: Duistermaat, 1996; Taylor, 1981; Hörmander, 1983]

Housekeeping chores

Proof of (ii): follows from

(iv) Basic estimate for oscillatory integrals: suppose that $\psi \in C^\infty(\mathbf{R}^n), \nabla\psi(\mathbf{x}_0) \neq \mathbf{0}$, $(\mathbf{x}_0, -\nabla\psi(\mathbf{x}_0)) \notin WF(u)$. Then for any $\chi \in C_0^\infty(\mathbf{R}^n)$ supported in small enough nbhd of $\mathbf{x}_0$, and any $N \in \mathbf{N}$,

$$\int dx\, e^{i\tau\psi(\mathbf{x})}\chi(\mathbf{x})u(\mathbf{x}) = O(\tau^{-N}),\ \tau \to \infty$$

Housekeeping chores

Proof of (iv): choose nbhd $X \times \Xi$ of $(\mathbf{x}_0, -\nabla\psi(\mathbf{x}_0))$ as in definition: conic, i.e. $(\mathbf{x}, \boldsymbol{\xi}) \in X \times \Xi \Rightarrow (\mathbf{x}, \tau\boldsymbol{\xi}) \in X \times \Xi$, $\tau > 0$.

Choose $a \in C^\infty(\mathbf{R}^n \setminus \{0\})$ homogeneous of degree 0 ($a(\boldsymbol{\xi}) = a(\boldsymbol{\xi}/|\boldsymbol{\xi}|)$) for $|\boldsymbol{\xi}| > 1$ so that $a(\boldsymbol{\xi}) = 0$ if $\boldsymbol{\xi} \notin \Xi$ or $|\boldsymbol{\xi}| \le 1/2$, $a(\boldsymbol{\xi}) = 1$ if $|\boldsymbol{\xi}| > 1$ and $\boldsymbol{\xi} \in \Xi_1 \subset \Xi$, another conic nbhd of $-\nabla\psi(\mathbf{x}_0)$.

Pick $\chi_1 \in C_0^\infty(\mathbf{R}^n)$ st $\chi_1 \equiv 1$ on $\mathrm{supp}\chi$, and write

$$\chi(x)u(x) = \chi_1(x)(2\pi)^{-n} \int d\boldsymbol{\xi}\, e^{i\mathbf{x}\cdot\boldsymbol{\xi}} \widehat{\chi u}(\boldsymbol{\xi})$$

Housekeeping chores

$$= \chi_1(x)(2\pi)^{-n}\int d\boldsymbol{\xi}\, e^{i\mathbf{x}\cdot\boldsymbol{\xi}} g_1(\boldsymbol{\xi})$$

$$+\chi_1(x)(2\pi)^{-n}\int d\boldsymbol{\xi}\, e^{i\mathbf{x}\cdot\boldsymbol{\xi}} g_2(\boldsymbol{\xi})$$

in which $g_1 = a\widehat{\chi u}$, $g_2 = (1-a)\widehat{\chi u}$

Housekeeping chores

So

$$\int dx\, e^{i\tau\psi(\mathbf{x})}\chi(\mathbf{x})u(\mathbf{x})$$

$$= \sum_{j=1,2} \int dx \int d\boldsymbol{\xi}\, e^{i(\tau\psi(x)+\mathbf{x}\cdot\boldsymbol{\xi})}\chi_1(x)g_j(\boldsymbol{\xi})$$

Housekeeping chores

For $\boldsymbol{\xi} \in \mathrm{supp}(1-a)$ (excludes a conic nbhd of $-\nabla\psi(\mathbf{x}_0)$), can write

$$e^{i(\tau\psi(\mathbf{x})+\mathbf{x}\cdot\boldsymbol{\xi})}$$

$$= [-i|\tau\nabla\psi(\mathbf{x})+\boldsymbol{\xi}|^{-2}(\tau\nabla\psi(\mathbf{x})+\boldsymbol{\xi})\cdot\nabla]^p e^{i(\tau\psi(\mathbf{x})+\mathbf{x}\cdot\boldsymbol{\xi})}$$

Housekeeping chores

Can guarantee that $|\tau\nabla\psi(\mathbf{x})+\boldsymbol{\xi}|>0$ by choosing $\mathrm{supp}\chi_1$ suff. small, so that in dom. of integration $\nabla\psi(\mathbf{x})$ is close to $\nabla\psi(\mathbf{x}_0)$. In fact, for $\boldsymbol{\xi}\in\mathrm{supp}(1-a)$, $\mathrm{supp}\chi_1$ small enough, and $\mathbf{x}\in\mathrm{supp}\chi_1$,

$$|\tau\nabla\psi(\mathbf{x})+\boldsymbol{\xi}|>C\tau$$

for some $C>0$. Exercise: prove this!

Housekeeping chores

Substitute and integrate by parts, use above estimate to get

$$\left| \int dx \int d\boldsymbol{\xi}\, e^{i(\tau\psi(\mathbf{x})+\mathbf{x}\cdot\boldsymbol{\xi})}\chi_1(\mathbf{x})g_2(\boldsymbol{\xi}) \right| \leq const.\tau^{-N}$$

for any N.

Note that for $\boldsymbol{\xi} \in \mathrm{supp} a$,

$$|\widehat{\chi u}(\boldsymbol{\xi})| \leq \mathrm{const.}|\boldsymbol{\xi}|^{-p}$$

for any p (with p-dep. const, of course!).

Housekeeping chores

Follows that

$$h(\mathbf{x}) = \int d\boldsymbol{\xi}\, e^{i\mathbf{x}\cdot\boldsymbol{\xi}} g_1(\boldsymbol{\xi})$$

converges absolutely, also after differentiating any number of times under the integral sign.

Housekeeping chores

therefore $h \in C^\infty(\mathbf{R}^n)$, whence

$$\int dx \int d\boldsymbol{\xi}\, e^{i(\tau\psi(\mathbf{x})+\mathbf{x}\cdot\boldsymbol{\xi})}\chi_1(\mathbf{x})g_1(\boldsymbol{\xi})$$

$$= \int dx\, e^{i\tau\psi(\mathbf{x})}\chi_1(\mathbf{x})h(\mathbf{x})$$

Housekeeping chores

with integrand supported as near as you like to $\mathbf{x}_0$. Since $\nabla\psi(\mathbf{x}_0) \neq 0$, same is true of $\mathrm{supp}\chi_1$ provided this is chosen small enough; now use

$$e^{i\tau\psi(\mathbf{x})} = \tau^{-p}(-i|\nabla\psi(\mathbf{x})|^{-2}\nabla\psi(\mathbf{x}) \cdot \nabla)^p e^{i\tau\psi(\mathbf{x})}$$

and integration by parts again to show that this term is also $O(\tau^{-N})$ any N.

Housekeeping chores

Proof of (ii), for u integrable (Exercise: formulate and prove similar statement for distributions)

Equivalent statement: suppose that $\Phi : U \rightarrow \mathbf{R}^n$ is a diffeomorphism on an open $U \subset \mathbf{R}^n$, $\mathrm{supp}\, u \subset \Phi(U)$, $\mathbf{x}_0 \in U$, $\mathbf{y}_0 = \Phi(\mathbf{x}_0)$, and $(\mathbf{y}_0, \eta_0) \notin WF(u)$.

Claim: then $(\mathbf{x}_0, \boldsymbol{\xi}_0) \notin WF(u \circ \Phi)$, where $\boldsymbol{\xi}_0 = D\Phi(\mathbf{x}_0)^T \eta_0$.

Housekeeping chores

Need to show that if $\chi \in C_0^\infty(\mathbf{R}^n), \mathbf{x}_0 \in \mathrm{supp}\chi$ and small enough, then $\widehat{\chi u \circ \Phi}(\tau \boldsymbol{\xi}) = O(\tau^{-N})$ any N for $\boldsymbol{\xi}$ conically near $\boldsymbol{\xi}_0$. From the change-of-variables formula

$$\widehat{\chi u \circ \Phi}(\tau \boldsymbol{\xi}) = \int dx \chi(\mathbf{x})(u \circ \Phi)(\mathbf{x}) e^{i\tau \mathbf{x}\cdot\boldsymbol{\xi}}$$

$$= \int dy \, (\chi \circ \Phi^{-1})(\mathbf{y}) u(\mathbf{y}) e^{i\tau \boldsymbol{\xi}\cdot\Phi^{-1}(\mathbf{y})} \det D(\Phi^{-1})(\mathbf{y})$$

Set $j = \chi \circ \Phi^{-1} det D(\Phi^{-1})$. Note: $j \in C_0^\infty(\mathbf{R}^n)$ supported in nbhd $\mathcal{V}$ of $\mathbf{y}_0$ if χ supported in $\Phi^{-1}(\mathcal{V})$.

Housekeeping chores

MVT: for $\mathbf{y}$ close enough to $\mathbf{y}_0$,

$$\Phi^{-1}(\mathbf{y}) = \mathbf{x}_0 + \int_0^1 d\sigma D\Phi^{-1}(\mathbf{y}_0 + \sigma(\mathbf{y} - \mathbf{y}_0))(\mathbf{y} - \mathbf{y}_0)$$

Insert in exponent to get

$$\widehat{\chi u \circ \Phi}(\tau\boldsymbol{\xi}) = e^{i\tau\mathbf{x}_0\cdot\boldsymbol{\xi}} \int dy\, j(\mathbf{y})u(\mathbf{y})e^{i\tau\psi_{\boldsymbol{\xi}}(\mathbf{y})}$$

where

$$\psi_{\boldsymbol{\xi}}(\mathbf{y}) = (\mathbf{y} - \mathbf{y}_0) \cdot \int_0^1 d\sigma D\Phi^{-1}(\mathbf{y}_0 + \sigma(\mathbf{y} - \mathbf{y}_0))^T\boldsymbol{\xi}$$

RICE

Housekeeping chores

Since

$$\nabla \psi_{\boldsymbol{\xi}}(\mathbf{y}_0) = D\Phi^{-1}(\mathbf{y}_0)\boldsymbol{\xi}$$

claim now follows from basic thm on oscillatory integrals.

Housekeeping chores

Proof of (iii): Function of compact supp, jumping across $\phi = 0$

$$u = \chi H(\phi)$$

with χ smooth, $H =$ Heaviside function ($H(t) = 1, t > 0$ and $H(t) = 0, t < 0$).

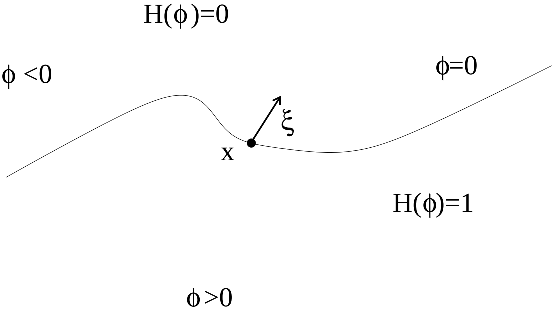
Pick $\mathbf{x}_0$ with $\phi(\mathbf{x}_0) = 0$. Surface $\phi = 0$ regular near $\mathbf{x}_0$ if $\nabla\phi(\mathbf{x}_0) \neq 0$ - assume this.

Housekeeping chores

Suffices to consider case of $\chi \in C_0^\infty(\mathbf{R}^n)$ of small support cont'g $\mathbf{x}_0$. Inverse Function Thm $\Rightarrow$ exists diffeo Φ mapping nbhd of $\mathbf{x}_0$ to nbhd of 0 so that $\Phi(\mathbf{x}_0) = 0$ and $\Phi_1(\mathbf{x}) = \phi(\mathbf{x})$. Fact (ii) $\Rightarrow$ reduce to case $\phi(\mathbf{x}) = x_1$ - Exercise: do this special case!

RICE

Wavefront set of a jump discontinuity



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

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).

Formalizing the reflector concept

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)

RICE

Microlocal property of differential operators

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

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

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

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

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:

Then $(\mathbf{x}_0, \xi_0) \notin WF(P(\mathbf{x}, D)u)$

That is, $WF(Pu) \subset WF(u)$.

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...

Proof

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

where $\phi_\alpha \in \mathcal{D}(X)$ is a linear combination of derivatives of ϕ and the a_αs. 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.**

2. Linearization, High frequency Asymptotics and Imaging

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:

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

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

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

Integral representation of linearized operator

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 for smooth v:

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

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

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

$$\tau(\mathbf{x}; \mathbf{x}_s) \sim \frac{r}{v(\mathbf{x}_s)},\ \ r = |\mathbf{x} - \mathbf{x}_s| \to 0$$

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

$$\nabla \cdot (a^2\nabla\tau) = 0;\ \ a \sim \frac{1}{4\pi r},\ r \to 0$$

Add geometric optics...

Why should this seem reasonable: formally, for constant v, G solves radiation problem for $w = \delta$:

$$G(\mathbf{x}, t; \mathbf{x}_s) = \frac{\delta\left(t - \frac{r}{v}\right)}{4\pi r}$$

so GO approx holds with $\tau(\mathbf{x}; \mathbf{x}_s) = |\mathbf{x} - \mathbf{x}_s|/v = r/v$ and $a = (4\pi r)^{-1}$ - in fact, it's not an approximation (R=0)!

Exercise: Verify that τ, a as given here, satisfy the eikonal and transport equation.

Add geometric optics...

Suppose

- v is const near $\mathbf{x} = \mathbf{x}_s$ (simplifying assumption - can be removed)
- τ smooth & satisfies eikonal equation for $r > 0$, $= r/v(\mathbf{x}_s)$ for small r
- a smooth & satisfies transport equation for $r > 0$, $= 1/4\pi r$ for small r

Then

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

is locally square-integrable

Add geometric optics...

(Hindsight!) Set

$$R_1(\mathbf{x}, t; \mathbf{x}_s) = \int_0^t ds R(\mathbf{x}, s; \mathbf{x}_s)$$

Will show that

$$R_1(\cdot, \cdot; \mathbf{x}_s) \in C^1(\mathbf{R}, L^2(\mathbf{R}^3)) \cap C^0(\mathbf{R}, H^1(\mathbf{R}^3))$$

which is sufficient.

Add geometric optics...

$$R_1(\mathbf{x}, t; \mathbf{x}_s) = \int_0^t ds G(\mathbf{x}, s; \mathbf{x}_s) - a(\mathbf{x}; \mathbf{x}_s) H(t - \tau(\mathbf{x}; \mathbf{x}_s))$$

Compute

$$\left(\frac{1}{v^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) R_1$$

Use calculus rules (why are these valid?). Expl:

$$\nabla a\delta(t-\tau) = (\nabla a)\delta(t-\tau) - a\nabla\tau\delta'(t-\tau)$$

(drop arguments for sake of space...)

Add geometric optics...

$$= \delta(\mathbf{x} - \mathbf{x}_s)H(t) - a\left(\frac{1}{v^2} - |\nabla\tau|^2\right)\delta'(t-\tau)$$

$$+(2\nabla\tau\cdot\nabla a + \nabla^2\tau a)\delta(t-\tau)$$

$$+\nabla^2 a H(t-\tau)$$

Add geometric optics...

Terms 2 & 3 vanish due to eikonal & transport -

$$= \delta(\mathbf{x} - \mathbf{x}_s)H(t) - \delta(\mathbf{x} - \mathbf{x}_s)H(t - \tau) + \text{ smooth}$$

$$= \text{smooth}$$

Quote Lions-Stolk result (++...) Q.E.D.

Add geometric optics...

Upshot: remainder R is more regular than the leading term - approximation of *leading singularity* or *high frequency asymptotics*

Local Geometric Optics

Main theorem of local geometric optics: if v is smooth in a nbhd of $\mathbf{x}_s$, then there exists a (possibly smaller) nbhd in which unique τ and a satisfying (a) and (b) exist, and are smooth except as indicated at $r = 0$.

Local Geometric Optics

Sketch of proof ("Hamilton-Jacobi theory"):

- basic ODE thm: solutions of IVP for Hamilton's Equations:

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

$$H(\mathbf{X}, \mathbf{\Xi}) = -\frac{1}{2}[1 - v^2(\mathbf{X})|\mathbf{\Xi}|^2]$$

$\mathbf{X}(0) = \mathbf{x}_s,\ v(\mathbf{x}_s)\mathbf{\Xi}(0) = \theta \in S^2$

- exponential polar coordinates: for $\mathbf{x}$ in nbhd of $\mathbf{x}_s$, exist unique $t, \mathbf{\Xi}(0)$ so that $\mathbf{X}(t) = \mathbf{x}$: set $\tau(\mathbf{x}) = t$

Local Geometric Optics

- for any trajectory $\mathbf{X}, \mathbf{\Xi}$ of HE, $t \mapsto H(\mathbf{X}(t), \mathbf{\Xi}(t))$ is constant; for these trajectories, IC $\Rightarrow |\mathbf{\Xi}(t)| = 1/v(\mathbf{X}(t))$
- $d\mathbf{X}/dt$ is parallel to $\nabla\tau$, in fact
- $\nabla\tau(\mathbf{X}(t)) = \mathbf{\Xi}(t) \Rightarrow$
- τ solves eikonal eqn

Exercise: complete this sketch to produce a proof - may assume v const near $\mathbf{x} = \mathbf{x}_s$

Local Geometric Optics

Hint: the 2nd step is crucial.

Idea: initial data is $(\mathbf{x}_s, \mathbf{\Xi}_0)$ where $\mathbf{\Xi}_0$ lies on sphere of radius $1/v(\mathbf{x}_s)$. Choose curve in sphere parameterized by $s \in \mathbf{R}$, passing through $\mathbf{\Xi}_0$ at $s = 0$;

develop ODE for

$$t \mapsto \left(\frac{\partial X}{\partial s}(t, \mathbf{\Xi}_0)\right)^T \frac{\partial X}{\partial t}(t, \mathbf{\Xi}_0)$$

init val $= 0 \Rightarrow$ always $= 0 \Rightarrow \mathbf{\Xi}$ perp to level surface of τ

Local Geometric Optics

Note: geometric optics *ray* $t \mapsto \mathbf{X}(t)$ is geodesic of Riemannian metric $v^{-2} \sum_{i=1}^{3} dx_i \otimes dx_i$

v smooth $\Rightarrow$ distance to nearest conjugate point > 0.

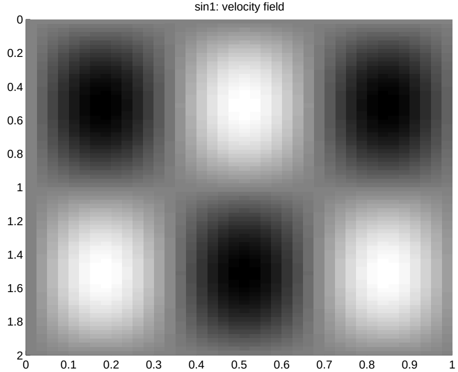
Numerics, and a caution

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

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).

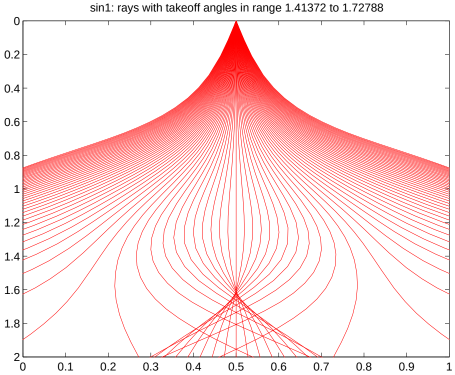
Upon caustic formation, the simple geometric optics field description above is no longer correct (Ludwig, 1966).

A caustic example (1)



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

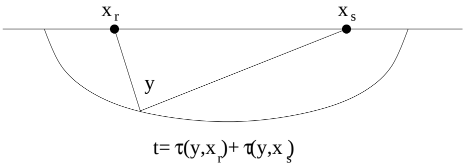
A caustic example (2)



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

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



The linearized operator as Generalized Radon Transform

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?].

Q: What does $\simeq$ mean?

A: It means "differs by something smoother".

In theory: develop R in series of terms of decreasing order of singularity

asymptotic: G - sum of N terms $\in C^{N-2}$

In practice, first term suffices (can formalize this with modification of wavefront set defn).

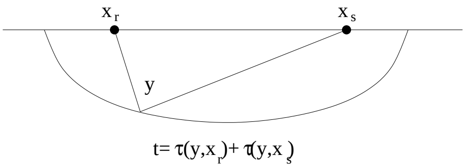
GRT = "Kirchhoff" modeling

$\mathrm{supp}\, r \subset$ simple geometric optics domain $\Rightarrow$

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

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

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 $\tau(\mathbf{x}_s, \mathbf{x}) = |\mathbf{x}_s - \mathbf{x}|/v$!



2. Linearization, High frequency Asymptotics and Imaging

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 v, r functions of $z = x_3$ only, all sources and receivers at $z = 0$

$\Rightarrow$ system is translation-invariant in x_1, x_2

$\Rightarrow$ Green's function G its perturbation δG, and the idealized data $\delta G|_{z=0}$ only functions of t and *half-offset* $h = |\mathbf{x}_s - \mathbf{x}_r|/2$.

Standard Processing

$\Rightarrow$ *only one seismic experiment*, equivalent to any *common midpoint gather* ("CMP").

This isn't really true - *look at the data!!!*

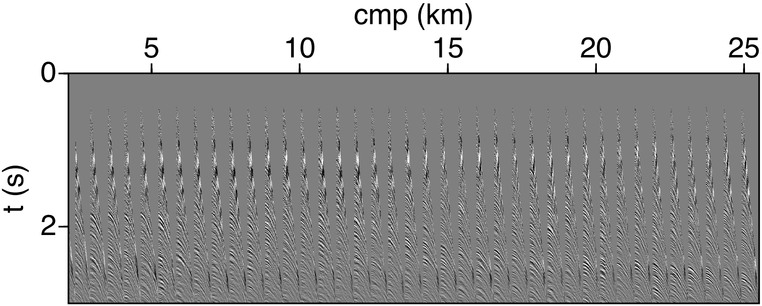
Standard Processing

Example: Mobil Viking Graben data

Released 1994 by Mobil R&D as part of workshop exercise ("invert this!")

North Sea "2D" data, i.e. single 25 km sail line, single 3 km streamer - passes near location of well, log shown in Part I

Standard Processing

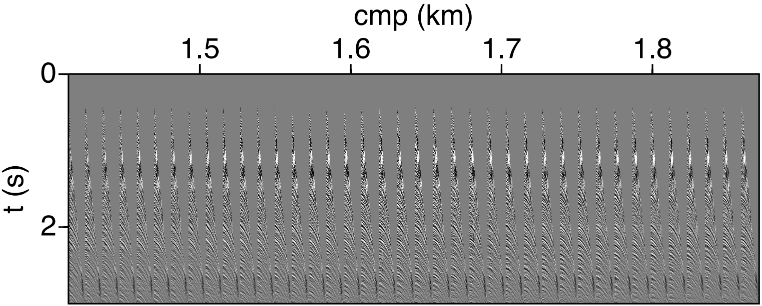


Sort to CMP gathers (common $\mathbf{x}_m = \mathbf{x}_s + \mathbf{x}_r/2$), extract every 50th - approx. 600 m between CMP locations

Standard Processing

However the "locally layered" idea 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



39 consecutive CMP gathers (1002-1040), distance between values of $\mathbf{x}_m = 12.5$ m

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

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

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

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

$$\times e^{i\omega(t - \tau(\mathbf{x}, x_r) - \tau(\mathbf{x}, x_s))}$$

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*.

Anatomy of an adjoint

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

$$= \int dt \int dh\, d(t,h) A(z(t,h),h) R(z(t,h))$$

$$= \int dz\, R(z) \int dh\, \frac{\partial t}{\partial z}(z,h) A(z,h) d(t(z,h),h)$$

$$= \int dz\, r(z) (F[v]^* d)(z)$$

RICE

Anatomy of an adjoint

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** $Sf(z) = \int dh\, f(z,h)$

RICE

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

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!

Particularly nice transformation: define $t_0 =$ *two-way vertical travel time* for z (depth):

$$t_0(z) = 2\int_0^z \frac{1}{v}$$

and its inverse function z_0

RMS (or NMO) velocity:

$$\bar{v}(t_0)^2 = \frac{1}{t_0}\int_0^{t_0} d\tau\, v(z_0(\tau))^2$$

Then ("Dix's formula") $\bar{t}(t_0, h) = t(z_0(t_0), h)$

$$= \sqrt{t_0^2 + 4h^2/\bar{v}^2(t_0)} + O(h^4)$$

which is exactly what the constant-v formula would be, if $\bar{v}$ were constant - *hyperbolic moveout*

Exercise: Prove this [hint: use eikonal, presumed symmetry of $t(z, h)$ to derive ODE for $\partial^2 t/\partial h^2(z, 0)$, solve]

NMO operator (as usually construed):

$$\bar{N}[\bar{v}]d(t_0, h) = d(\bar{t}(t_0, h), h)$$

Now make everything dependent on $\mathbf{x}_m$ and you've got standard processing.

[LIVE DEMO - Mobil AVO data, Seismic Unix]

An interesting observation: if $d(t, h)$ conforms to the layered etc. etc. approximation, i.e.

$$d(t, h) = F[v]r(t, h)$$

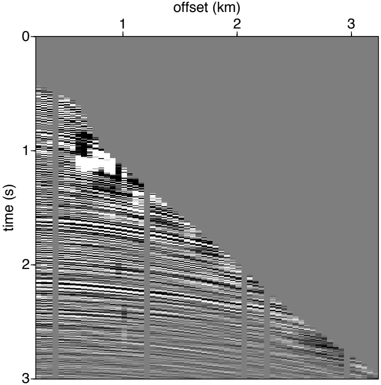
then

$$N[v]d(z, h) = d(t(z, h), h) = (\text{amplitude factor } \times r(z)$$

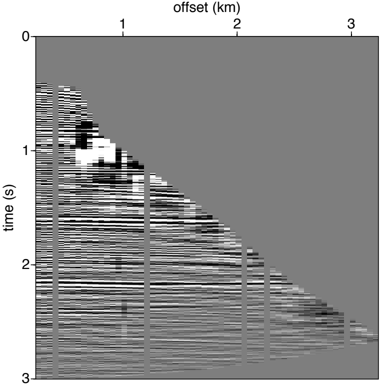
i.e. except for the amplitude factor, this part of $F[v]^*$ produces function independent of h - amplitude is smooth, r is oscillatory, should be obvious

Similar if use t_0 as depth variable instead of z

Example: apply NMO operator $N[v]$ to CMP 1040 from Mobil AVO data:

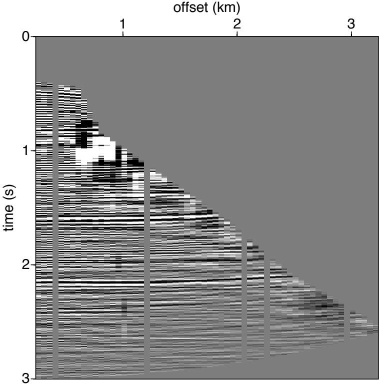


CMP 1040

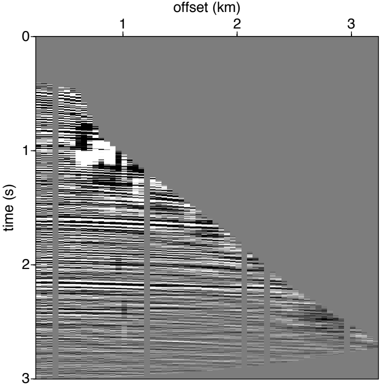


NMO: v=1.5, 1.8, 2.2, 2.4

RICE



NMO: v=1.5, 1.7, 2.0, 2.2



NMO: v=1.5, 2.0, 2.4, 2.6

Upshot: if v (or $\bar{v}$) chosen "well" (matching trend of velo in earth?), then NMO output is mostly indep of $h = \mathit{flat}$

$\Rightarrow$ method for determining v (and r) - *velocity analysis*

Sounds like voodoo - what does it have to do with inversion?

Stay tuned!

2. Linearization, High frequency Asymptotics and Imaging

Multioffset ("Prestack") Inversion, après Beylkin

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.*

Multioffset ("Prestack") Inversion, après Beylkin

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

- $WF(F_\delta[v]^* F_\delta[v] r) \subset WF(r)$
- if $S^{\rm obs} = S[v] + F_\delta[v] r$ (data consistent with linearized model), then $F_\delta[v]^*(S^{\rm obs} - S[v])$ is an image of r
- an operator $F_\delta[v]^\dagger$ exists for which $F_\delta[v]^\dagger(S^{\rm obs} - S[v]) - r$ 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"):

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

Outline of proof

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

symbol $p \in C^\infty$: for some $m \in \mathbf{R}$, all multiindices α, β, and all compact $K \subset \mathbf{R}^n$, there exist $C_{\alpha,\beta,K} \geq 0$ for which

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

order of p is inf of all such m (or $-\infty$ if there is none)

Outline of proof

Explicit computation of **symbol** p of $F[v]^*F[v]$ in terms of rays, amplitudes - for details, see WWS: Math Foundations.

[Symbol in terms of operator ($m =$ order): for $\phi \in C_0^\infty(\mathbf{R}^n)$

$$p(\mathbf{x}, \boldsymbol{\xi})\phi(\mathbf{x}) = e^{i\mathbf{x}\cdot\boldsymbol{\xi}} p(\mathbf{x}, D) e^{-i\mathbf{x}\cdot\boldsymbol{\xi}} \phi(\mathbf{x}) + O(|\boldsymbol{\xi}|^{m-1})$$

- will return to this fact!]

Microlocal Property of 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; imaging property of prestack Kirchhoff migration follows.

Microlocal Property of ΨDOs

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 (much 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 $\phi \in C_0^\infty(X')$, estimate $\widehat{\phi P(\mathbf{x}, D)}u$. 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.

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

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

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

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

$$\times e^{i(\mathbf{x}-\mathbf{y})\cdot\boldsymbol{\xi}} u(\mathbf{y})$$

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,

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

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:

$$\widehat{\phi P(\mathbf{x}, D)}(\psi u)(\tau\eta)$$

$$= \int dx \int d\xi \, P(\mathbf{x}, \boldsymbol{\xi})\phi(\mathbf{x})\hat{\psi u}(\xi)$$

$$\times e^{i\mathbf{x}\cdot(\boldsymbol{\xi}-\tau\eta)}$$

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 $|\theta - \eta| \le \epsilon \Rightarrow \theta \in \Xi$, per the defn of Ξ'. Since $X \times \Xi$ is disjoint from the wavefront set of u, for a sequence of constants C_N, $|\hat{\psi u}(\tau\theta)| \le C_N \tau^{-N}$ 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 migration works, at least in the simple geometric optics regime.

An Example

In what sense can this work with "bandlimited" ($w \neq \delta$) data?

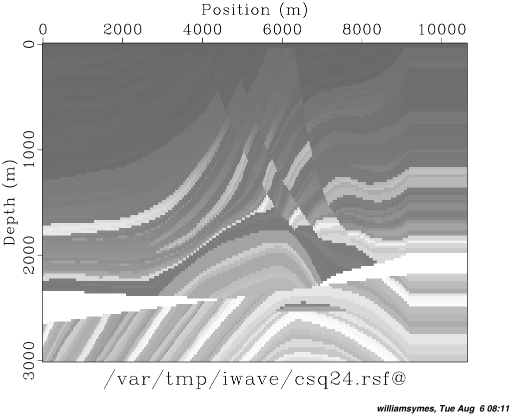
$F[v]^*F[v]r$ then does not have any singularities, even if r does, so no wave front set.

Answer: "ghost of departed wavefront set": as $w \rightarrow \delta$, $F[v]^*F[v]r \rightarrow$ a distribution with wavefront set $\subset WF(r)$.

RICE

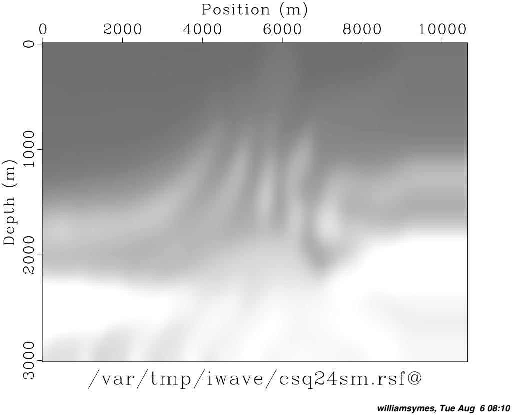
An Example

Marmousi c^2



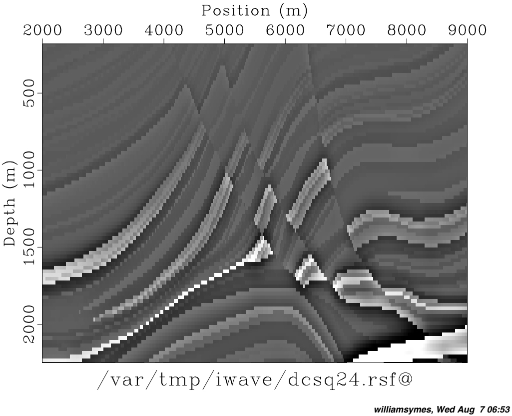
An Example

Marmousi v^2



An Example

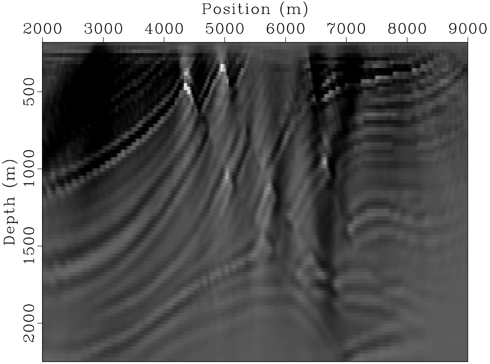
Marmousi $\delta(c^2) = 2vr$



RICE

An Example

Marmousi $F[v]^*F[v]r$



acdadj_line100m_mute.rsf@

williamsymes, Tue Aug 6 08:42

Symbol and Spectrum

Recall that for $p(x, D)$ of order m, $\phi \in C_0^\infty(\mathbf{R}^n)$,

$$p(\mathbf{x}, \boldsymbol{\xi})\phi(\mathbf{x}) = e^{i\mathbf{x}\cdot\boldsymbol{\xi}} p(\mathbf{x}, D) e^{-i\mathbf{x}\cdot\boldsymbol{\xi}} \phi(\mathbf{x}) + O(|\boldsymbol{\xi}|^{m-1})$$

Exercise: give a proof in case $p(x, D)$ is differential op of order m

Double Bonus Exercise: give a proof

Symbol and Spectrum

Exercise: Write an application using Pysit to approximate the symbol of $F[v]^*F[v]$ at $(\mathbf{x}, \boldsymbol{\xi}) \in T^*X$

Symbol and Spectrum

Special class of symbols: those with asymptotic expansions

$$p(\mathbf{x},\boldsymbol{\xi}) = \sum_{l \in \mathbf{N}} p_{m-l}(\mathbf{x},\boldsymbol{\xi})$$

in which p_k is a symbol, positively homogeneous in $\boldsymbol{\xi}$ of order k.

Consequence of a theorem of Borel: any such asymptotic series defines a symbol

Symbol and Spectrum

Fact: $F[v]^*F[v]$ is a ΨDO whose symbol has an asymptotic expansion (assuming simple ray geometry)

Principal symbol $=$ leading order term p_m

p is *microlocally elliptic* in a open conic nbhd Γ of $(\mathbf{x}_0, \boldsymbol{\xi}_0)$ if $p_m \neq 0$ in Γ: in any closed subnbhd $\Gamma_0 \subset \Gamma$, there is $K > 0$ so that for $(\mathbf{x}, \xi) \in \Gamma_0$,

$$|p_m(\mathbf{x}, \xi)| \geq K|\xi|^m$$

Symbol and Spectrum

Assume p microlocally elliptic at $(\mathbf{x}_0, \boldsymbol{\xi}_0)$, $\phi \in C_0^\infty(\mathbf{R}^d)$ supported near $\mathbf{x}_0$ - then

$$p(\mathbf{x}, D)e^{-i\mathbf{x}\cdot\boldsymbol{\xi}_0}\phi(\mathbf{x}) = p_m(\mathbf{x}, \boldsymbol{\xi}_0)e^{-i\mathbf{x}\cdot\boldsymbol{\xi}_0}\phi(\mathbf{x}) + O(|\boldsymbol{\xi}_0|^{m-1})$$

and remainder is rel. small for large $\boldsymbol{\xi}_0 \Rightarrow$ localized oscillatory "approximate eigenfunction"

Much more precise results available (eg. Demanet-Ying) - connect principal symbol to spectra of operators defined by ΨDO

Asymptotic Prestack Inversion

Recall: in layered case,

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

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

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

In particular, the normal operator $F[v]^*F[v]$ is an elliptic PDO.

$\Rightarrow$ normal operator is *asymptotically invertible*

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 prestack simple geometric optics case: due to Beylkin (1985).

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

Inversion aperture

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

$WF(r) \subset \Gamma[v] \Rightarrow WF(F[v]^*F[v]r) = WF(r)$

$\Rightarrow F[v]^*F[v]$ "acts invertible"

$(\mathbf{x}, \boldsymbol{\xi}) \in \Gamma[v] \Leftrightarrow F[v]^*F[v]$ microlocally elliptic at $(\mathbf{x}, \boldsymbol{\xi})$

Ray-geometric construction of $\Gamma[v]$ - later!

Inversion aperture

Beylkin: with proper choice of amplitude $b(\mathbf{x}_r, t; \mathbf{x}_s)$, the integral operator (modification of the integral representation of F^*)

$$F[v]^\dagger d(\mathbf{x}) =$$

$$\int\int\int dx_r\, dx_s\, dt\, b(\mathbf{x}_r, t; \mathbf{x}_s)\delta(t - \tau(\mathbf{x}; \mathbf{x}_s) - \tau(\mathbf{x}; \mathbf{x}_r))$$

$$\times d(\mathbf{x}_r, t; \mathbf{x}_s)$$

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.

aliases for numerical implementation: Generalized Radon Transform ("GRT") inversion, Ray-Born inversion, migration/inversion, true amplitude migration,...

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

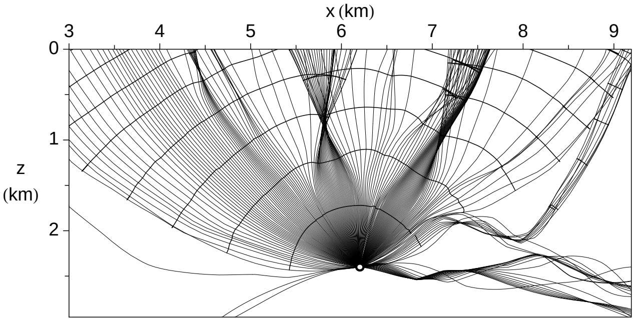
Apparent limitation: construction relies on simple geometric optics (no multipathing) - how much of this can be rescued?

An Example, cont'd

Apparently, quite a bit.

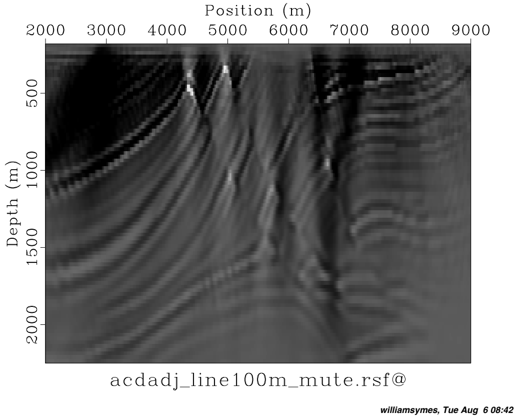
Marmousi (even smoothed v) generates many conjugate points, multipaths, caustics...

An Example, cont'd



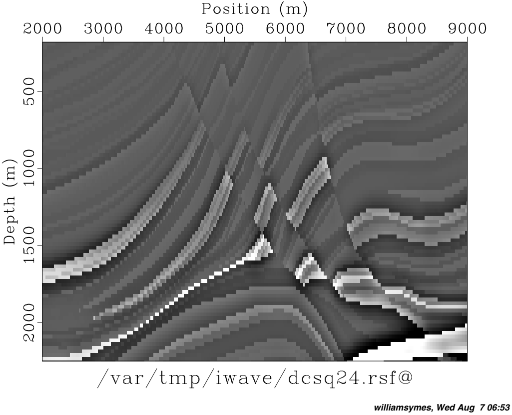
An Example, cont'd

yet $F[v]^*F[v]r$ is a good "image"...



An Example, cont'd

of r



An Example, cont'd

Of course $F[v]^* F[v] r$ just an "image"

Computation of $F[v]^\dagger F[v] r$ - not necessarily by integral representation - should restore amplitudes

An Example, cont'd

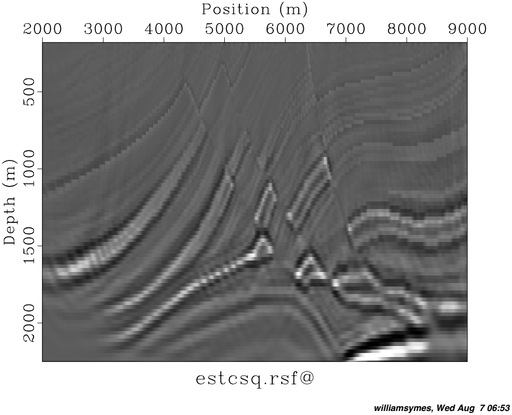
Inversion by iterative solution of

$$\min_r \|F[v]r - (d - \mathcal{F}[v])\|^2$$

- 60 shots, 10 Hz Ricker; 96 receivers 25 m spacing (classic IFP geometry, subsampled)
- 2-4 FD scheme, 24 m grid
- 50 conjugate gradient iterations
- reduces obj fcn to 20% of its initial value ($\|d - \mathcal{F}[v]\|^2$)

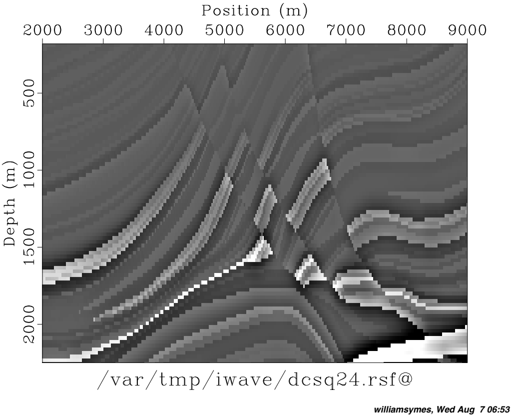
An Example, cont'd

reasonably good recovery...

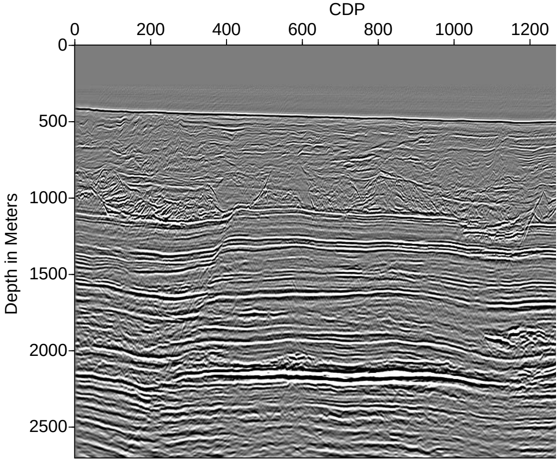


An Example, cont'd

of r (same grey scale!)



RICE



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.