Cross-scale Interaction in Seismic Inverse Scattering

William W. Symes

IPAM, February 2005

Seismic Inverse Scattering

- sample of seismic waves, acquired at Earth's surface, encodes Earth structure
- inverse scattering problem: decode it! [inverse problem = data assimilation / parameter estimation]
- primary subsurface exploration technique for petroleum industry, also used in civil & env. engineering, academic Earth Science
- dynamics of seismic waves are *linear*, but relation between data, dynamical parameters is *nonlinear* - so inverse problem is *nonlinear*
- key concepts underlying inversion methods in widespread use: linearization, adjoint state, ray theory, separation of scales
- critical need: better integration of nonlinearity
- primary industry variant uses active source; 95%+ of industry data acquired at sea

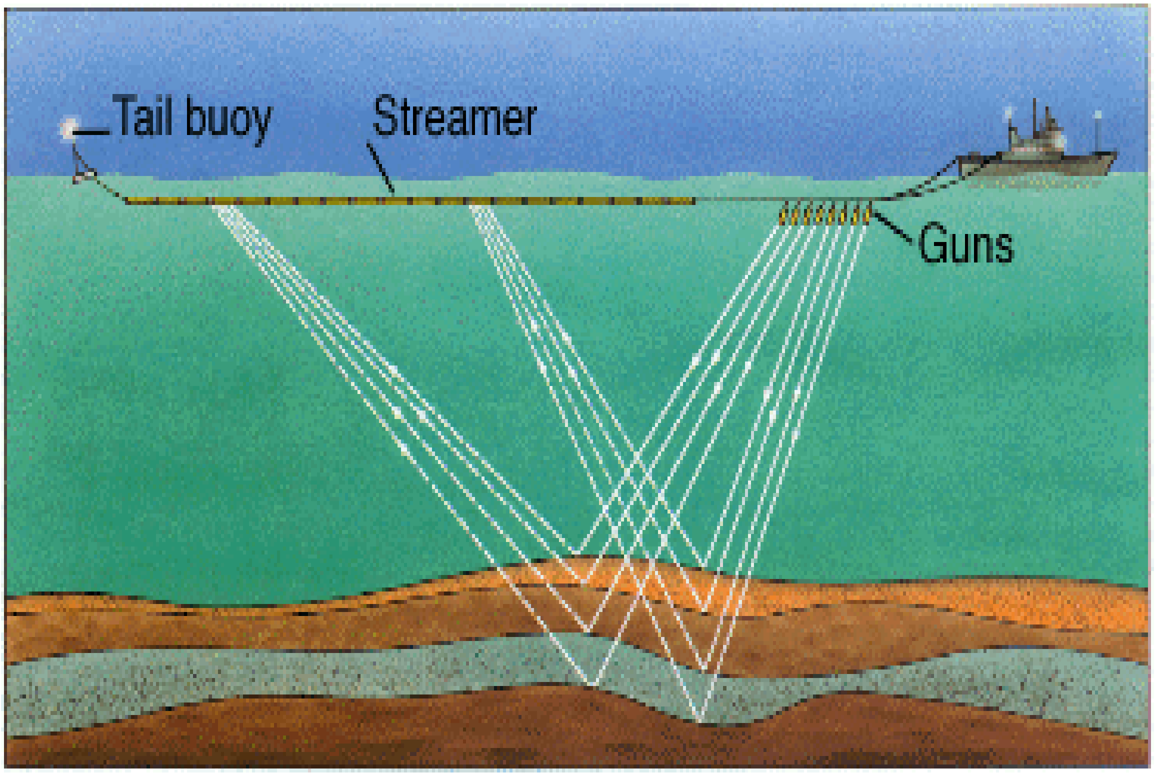
Outline

1. The Seismic Experiment
2. Length and Time Scales
3. State of the Art: Asymptotic Scale Separation and Partial Linearization
4. The Nonlinear Bit
5. Beyond Linearization

1. The Seismic Experiment

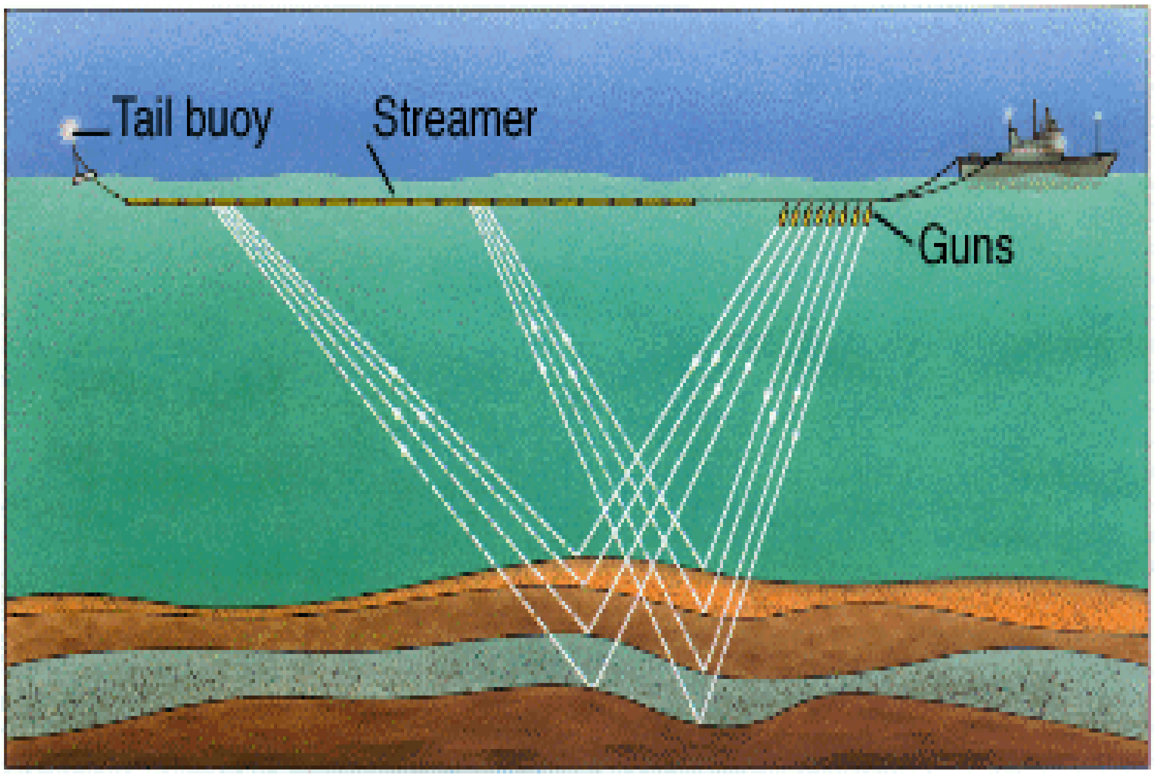
Marine reflection seismology, as practiced by the petroleum industry

Marine Reflection Seismology



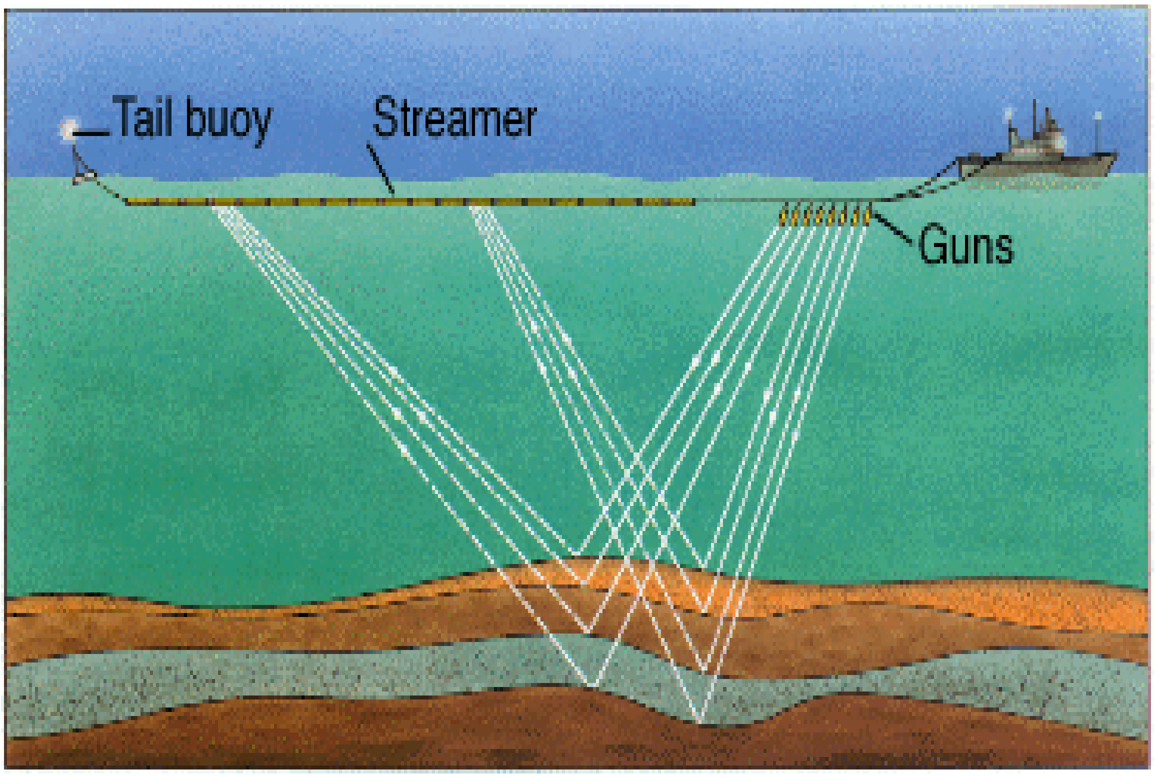
Ship tows *source* (compressed air gun) and *receivers* (hydrophone streamer). Typically records 100's of channels - recently 1000's. [thanks: Schlumberger]

Marine Reflection Seismology



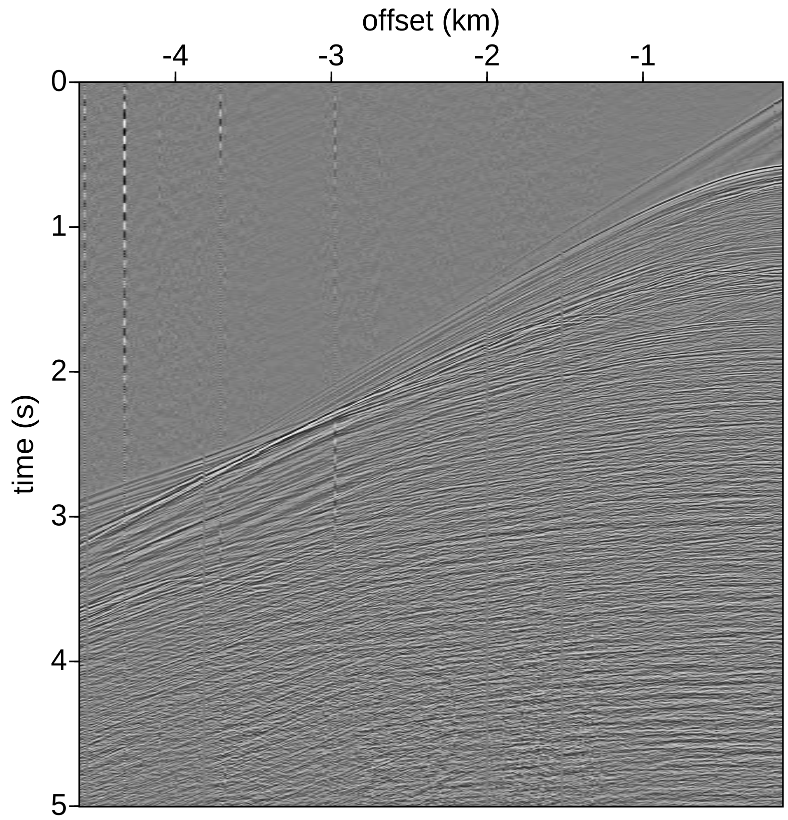
Typical distances: streamer length = 3 - 8 km, source spacing $\simeq$ hydrophone spacing = 10 - 25 m, parallel *sail lines* spaced 25 - 150 m apart.

Marine Reflection Seismology



Data acquisition - 24 $\times$ 7, source repetition every $\simeq$ 10 s. Typical modern survey covers 10 km $\times$ 10 km, produces data volume $\simeq$ 1 Tbyte,

Typical Marine Record



Shot record, Gulf of Mexico, Mississippi Canyon ca. 1985 (thanks: Exxon)

2. Length and Time Scales

Both seismic wavefield (dynamic state of system, sampled) and Earth structure (system control, to be estimated) are *broadband* - features at many scales - but the Earth is broader band!

Length/time Scales of Wavefield

Typical bandwidth of recorded data: 5 - 50 Hz $\simeq$ 3 octaves - limited by (1) response characteristics of recording equipment on both high and low end, (2) feasibility of energy input at low end, (3) absorption of high frequency components during propagation.

Recent improvements in acquisition equipment have pushed the limits, but fundamental physics $\Rightarrow$ a few octaves.

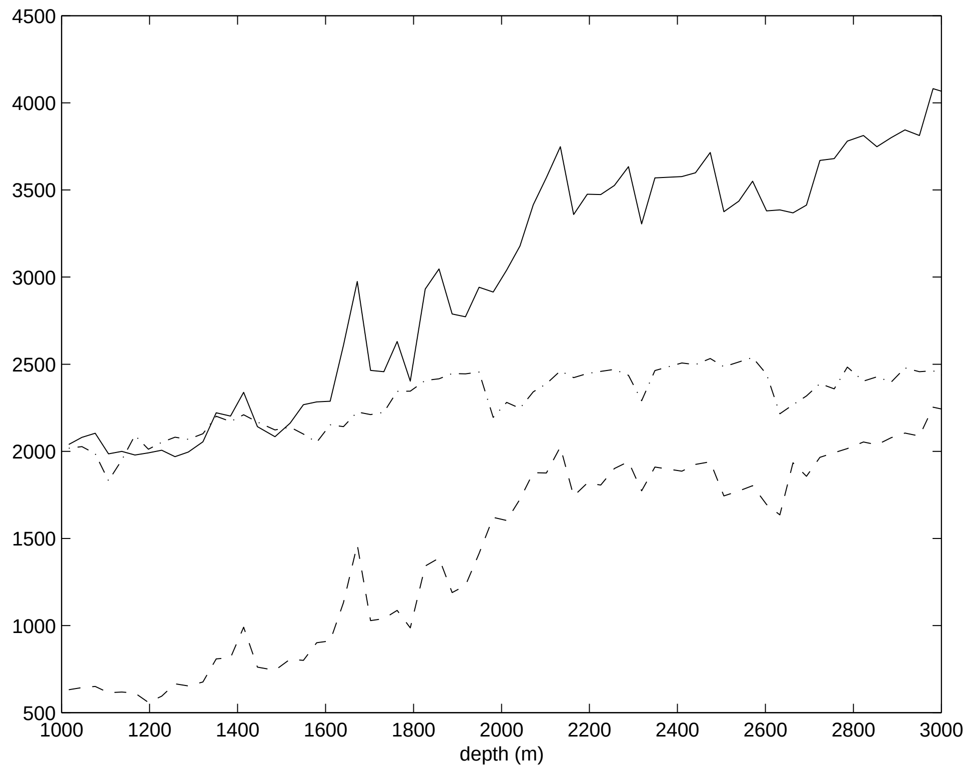
Periods of recorded waves: 20-200 ms. Typical velocities of waves in sedimentary rocks $\simeq$ 0.5 m/s (very near surface) - 5 m/ms (very hard sediments such as salts, anhydrites). So wavelenth range $\simeq$ 10 - 1000 m.

Typical wavelength of (compressional) wave in sands, shales $\simeq$ 100 m.

Mechanical Characteristics of Sedimentary Rock

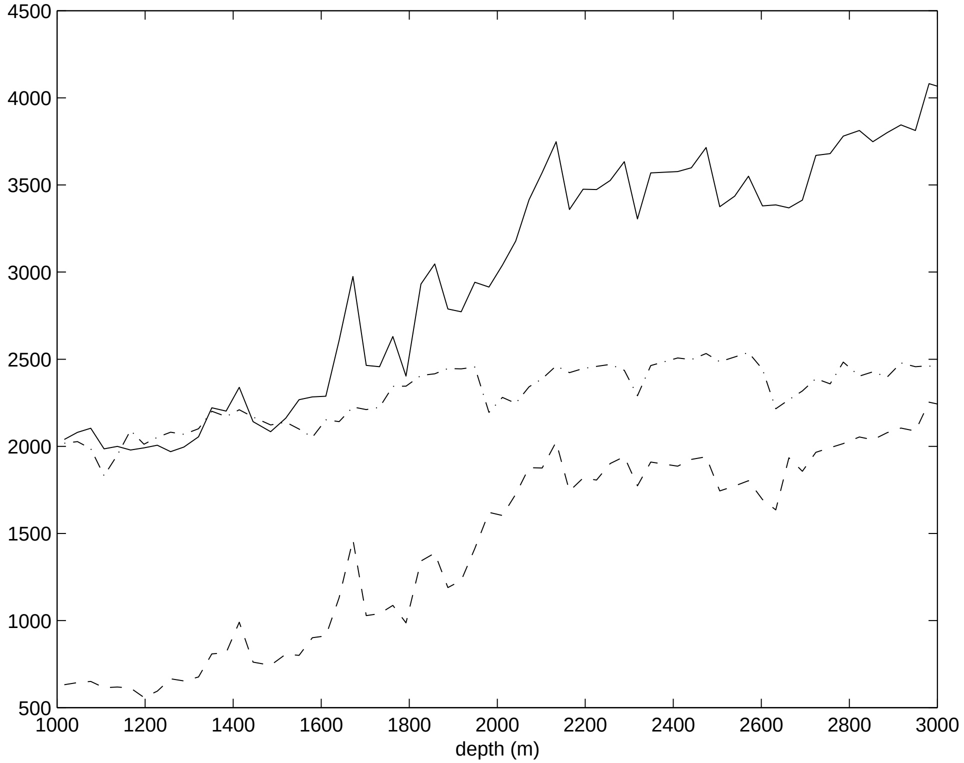
- to good approximation, sedimentary rocks interact *elastically* with seismic waves, and interaction is *linear* (amplitudes are small)
- to less good approximation, elasticity tensor is *isotropic*
- natural physical parameters of linear isotropic elastodynamics: *density* ρ, *compressional wave speed* v_p, and *shear wave speed* v_s (functions of position $\mathbf{x}$)
- (more or less) direct measurement of these quantities in boreholes: well logging - produces 1D cross section

Mechanical Characteristics of Sedimentary Rock



Well logs from North Sea borehole. Top: v_p (m/s); middle: ρ (kg/m^3); bottom: v_s (m/s). (thanks: Mobil R&D, Viking Graben). Original sampling: 0.25 m; subjected to 30 m moving average.

Mechanical Characteristics of Sedimentary Rock



Note (1) large variance at long (km’s) and short (10’s of m) scales (also at shorter scales, in original data); (2) relatively small variance of density

Scale Interactions

Interaction of seismic waves at various wavelengths with fluctuations in Earth mechanical properties at various scales, asymptotic simplifications:

- wavelength $\simeq$ correlation length: *scattering*, reflection, resonant regime interaction - no simple description except in limit of small amplitude fluctuations (single or "Born" scattering);
- wavelength $<<$ correlation length: *refraction*, asymptotics = geometric optics, ray theory
- wavelength $>>$ correlation length: *averaging*, asymptotics = homogenization, effective medium theory (eff. models tend to be same, with eff. params $\Rightarrow$ ignore!)

Critical observation: *data lacks long scales (km's) governing refraction*

Constant Density Acoustic Model

Notation for Data parameters: time t, source location $\mathbf{x}_s$, and receiver location $\mathbf{x}_r$, (vector) *half offset* $\mathbf{h} = \frac{\mathbf{x}_r - \mathbf{x}_s}{2}$, scalar half offset $h = |\mathbf{h}|$. Experiment = *shot*, single experiment data = *shot record*.

acoustic potential $u(\mathbf{x}, t)$, *sound velocity* $c(\mathbf{x})$ related to pressure p and particle velocity $\mathbf{v}$ by

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

Second order wave equation for potential

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

plus initial, boundary conditions. RHS models localized energy source. *Source wavelet* $w(t)$ determines frequency content of solution.

Inverse Scattering as Least Squares

Forward map: $\mathcal{F}[c] \equiv p|_Y$, where $Y = \{(t, \mathbf{x}_r, \mathbf{x}_s) : 0 \leq t \leq T, ...\}$ is *acquisition manifold*.

Inverse problem: given $d \in L^2(Y)$ find $c \in C$ s. t. $\mathcal{F}[c] \simeq d$.

Least squares formulation:

$$\min_{c \in C} \|\mathcal{F}[c] - d\|^2$$

An interesting question: What is a good choice of C? Would (1) honor actual complexity of Earth structure, (2) permit mathematically precise expression of physical insights on cross-scale interaction.

$\Rightarrow$ Research challenge!

Cross-scale Interaction and Regularity of $\mathcal{F}$

- (Stolk, 2000): with C = open set in L^∞, $\mathcal{F}$ *differentiable with loss of one derivative*, i.e. if $w \in H^s_{\rm loc}$ then $\mathcal{F}[c](\mathbf{x}_s, \mathbf{x}_r, \cdot) \in H^s_{\rm loc}$ (3D) but generally $D\mathcal{F}[c]\delta c(\mathbf{x}_s, \mathbf{x}_r, \cdot) \in H^{s-1}_{\rm loc}$ (sharp)
- Geom. Optics: δc *smooth* $\Rightarrow$ h.f. phases in $\mathcal{F}[c + \delta c]$ *shifted* rel $\mathcal{F}[c]$.
- Recall: data is *oscillatory* on km scale, lacks low frequencies
- Therefore δc is *smooth* $\Rightarrow \mathcal{F}[c]$ and $\mathcal{F}[c + \delta c]$ tend to be *nearly orthogonal* even when δc is small

Consequences for LS Inversion

- least squares function tends to *saturate*, i.e. remain near max, except when c is "right at long scales";
- fluctuations in angle between $\mathcal{F}[c]$, $\mathcal{F}[c+\delta c]$ as δc varies $\Rightarrow$ stationary points far from global min, *even when data is free of noise* ($d = \mathcal{F}[c]$)!!!
- Problems are so large that iterative methods (variants of Newton) are only feasible approach (3D: millions of unknowns, billions of equations) $\Rightarrow$ can only find stationary points.

Upshot: LS has had *little practical impact*.

3. State of the Art: Asymptotic Scale Separation and Partial Linearization

- Linearization (single scattering model) simplifies description of intrascale interaction - often can either ignore deviation from single scattering, or fake the data
- Scale separation: long scales in reference Earth model, short scales in perturbation $\Rightarrow$
  - minimizes linearization error
  - simplifies solution of LS formulation for short scales

(Partly) linearized inverse scattering

Formally, $c = v[1+r]$, $\mathcal{F}[v(1+r)] \simeq \mathcal{F}[v] + F[v]r$ where $F[v]r$ "=" $D\mathcal{F}[v](rv)$ is *linearized forward map* defined by

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

- *linearization error* $\mathcal{F}[v(1+r)] - \mathcal{F}[v] - F[v]r$ appears to be *smallest* when (1) v includes all long-scale features in model, (2) r contains only short-scale (oscillatory) features.
- "physical" reason for this: geom optics suggests no perturbation of long-scale velocity components $\Rightarrow$ no phase shift of short-scale wavefield components [no rigorous mathematical distillation of this observation currently known - research opportunity!]
- *linearized* inverse problem - given d, v, find r so that $F[v]r \simeq d$

Linearized inverse scattering

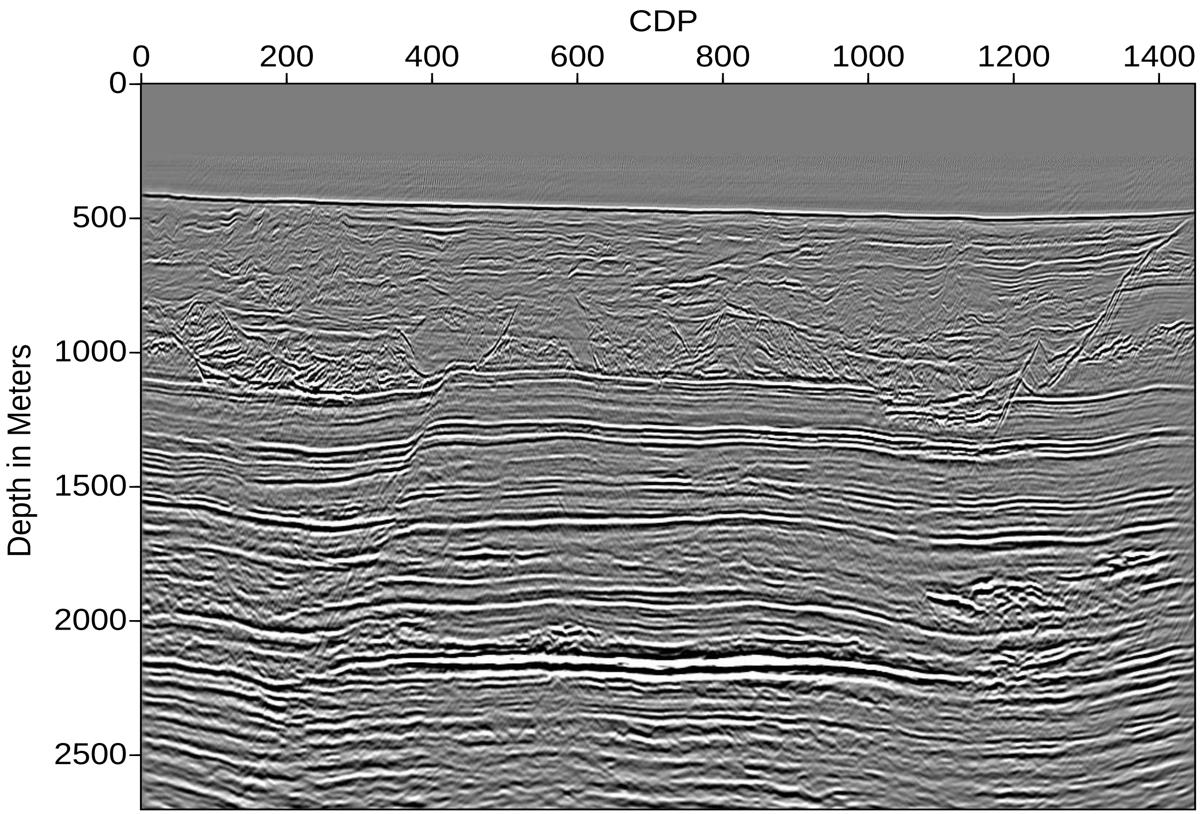
Study based on modern geom. optics ("microlocal analysis") by Beylkin 1985, Rakesh 1988, Nolan 1997, ten Kroode et al. 1998, Stolk 2000, others:

- At a stationary point r^*, $F[v]^*F[v]r^* = F[v]^*(d - \mathcal{F}[v])$
- Critical player: normal operator $F[v]^*F[v]$ is *pseudodifferential* (generically) and *microlocally elliptic*
- Pseudodifferential operators do not move short scale components
- $\Rightarrow r^*$ has same *locations of short scale components*, i.e. structure, as *migration of data* $F[v]^*d$
- $\Rightarrow$ for *image of structure*, don't need inversion (stat. point), migration = application of adjoint will do! ["convergence" of gradient optimization in one iteration!]

Linearized inverse scattering

- Computation of $F[v]^*$ ("migration operator"): adjoint state method ("wave equation migration"), usually coupled with one-way approximation, or direct use of asymptotic Green's function ("Kirchhhoff migration").
- $F[v]^*F[v]$ is *microlocally elliptic* $\Rightarrow$ oscillatory integral formulae approximating pseudoinverse for $F[v]$ ("asymptotic inversion")
- with Gaussian noise model, $r^* = (F[v]^*C_d^{-1}F[v] + C_m^{-1})^{-1}F[v]^*C_d^{-1}(d - \mathcal{F}[v])$ - if data, model stats are iid, then once again r^* has short scale components (structure) in same place as $F[v]^*d$ - so nothing about *structure* is gained unless nontrivial long-range correlations built into stats.
- Image structure = location of large short scale components = nonlinear functional of model - to some extent independent of (unknown) noise models.
- Amplitudes (actual values of r^*) strongly depend on noise model, but even more strongly on neglected physics (nonlinearity, anelasticity,...)!

$$\min_r \|F[v]r - (d - \mathcal{F}[v])\|^2, \text{ given "good" } v$$



Approximate linear least squares solution après Beylkin ("GRT inversion"), Mississippi Canyon, Gulf of Mexico, 2D survey (750 MB, 500 shots). Thanks: Exxon.

4. The Nonlinear Bit

Estimating the reference (long scale) velocity v is the nonlinear part of the partially linearized inverse scattering problem.

How it's done: *Extended Modeling*

Estimating v - Extended Models

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

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

commutes, i.e.

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

Extension is "invertible" iff $\bar{F}[v]$ has a *right parametrix* $\bar{G}[v]$, i.e. $I - \bar{F}[v]\bar{G}[v]$is smoothing, or more generally if $\bar{F}[v]\bar{G}[v]$ is pseudodifferential ("inverse except for wrong amplitudes"). Also require existence of a left inverse η for χ: $\eta\chi = \mathrm{id}$.

NB: The trivial extension - $\bar{X} = X, \bar{F} = F$ - is virtually never invertible.

Grand Example

The linearized model, rewritten with the aid of a Green's function $g = g[v]$ (and $w = \delta$ for simplicity):

$$F[v]r(\mathbf{x}_r, \mathbf{x}_s, t) = \frac{\partial^2}{\partial t^2} \int dx \int d\tau \, g(\mathbf{x}, \mathbf{x}_r, t - \tau) g(\mathbf{x}, \mathbf{x}_s, \tau) \frac{r(\mathbf{x})}{v^2(\mathbf{x})}$$

The Standard Extended Model: $\bar{X} = X \times H$, $H =$ offset range.

$$\bar{F}[v]\bar{r}(\mathbf{x}_r, \mathbf{x}_s, t) = \frac{\partial^2}{\partial t^2} \int dx \int d\tau \, g(\mathbf{x}, \mathbf{x}_r, t - \tau) g(\mathbf{x}, \mathbf{x}_s, \tau) \frac{2\bar{r}(\mathbf{x}, (\mathbf{x}_r - \mathbf{x}_s)/2)}{v^2(\mathbf{x})}$$

$\chi r(\mathbf{x}, \mathbf{h}) = r(\mathbf{x})$, $\eta \bar{r}(\mathbf{x}) = \frac{1}{|H|} \int_H dh \, \bar{r}(\mathbf{x}, \mathbf{h})$ ("stack").

$\bar{r} \in$ range of $\chi \Leftrightarrow$ plots of $\bar{r}(\cdot, \cdot, z, \mathbf{h})$ ("(prestack) image gathers") appear *flat*.

Reformulation of inverse problem

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

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

Proof: Hypothesis means

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

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

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

Q. E. D.

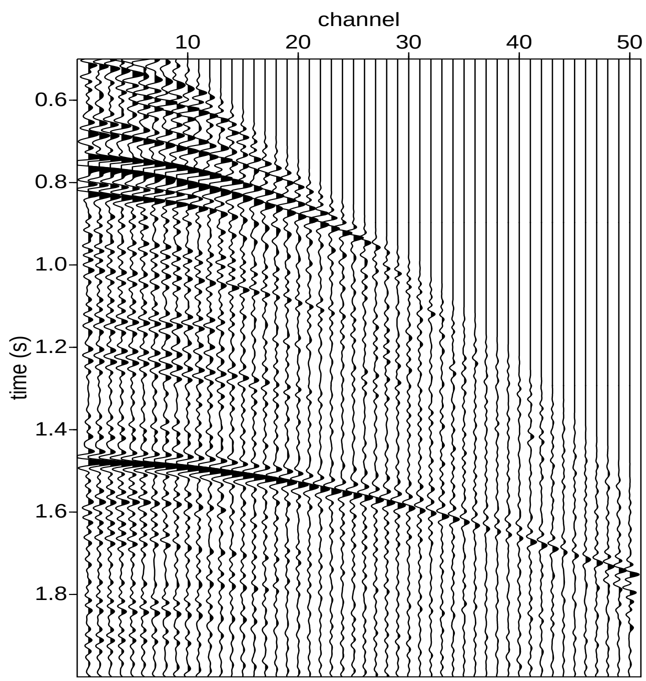
Application: Migration Velocity Analysis

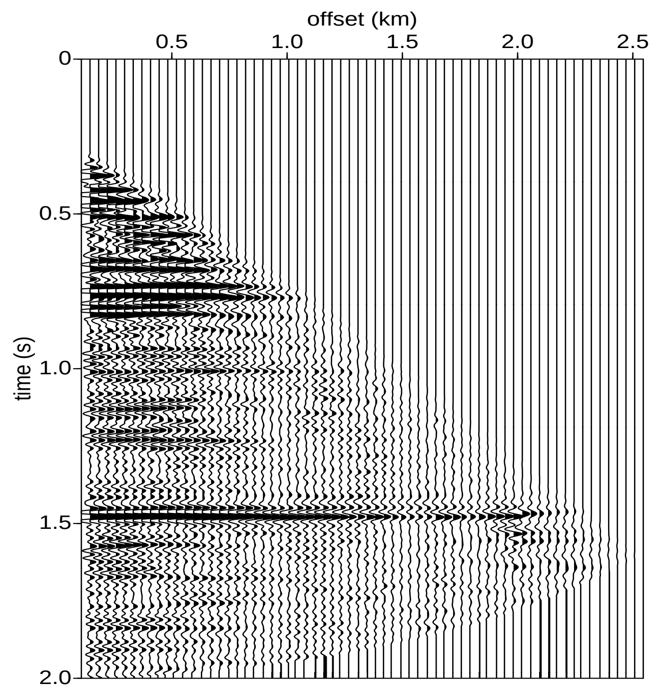
Membership in range of χ is *visually evident*

For the Standard Extended Model, $\bar{G}[v]d \in \mathcal{R}(\chi) \Leftrightarrow$ independent of $\mathbf{h}$.

$\Rightarrow$ industrial practice: adjust parameters of v *by hand* (!) until $\bar{G}[v]d$ exhibits visual characteristics of $\mathcal{R}(\chi)$ - “flatten the image gathers”.

Practically: insist only that $\bar{F}[v]\bar{G}[v]$ be pseudodifferential, so adjust v until $\bar{G}[v]d$ is “smooth” in $\mathbf{h}$.





,

Left: shot record (d) from North Sea survey (thanks: Shell Research), lightly pre-processed.
Right: restriction of $\bar{G}[v]d^{\mathrm{obs}}$ to $x, y = \mathrm{const}$ (function of depth, offset): shows relative smoothness in h (offset) for properly chosen v.

5. Beyond Linearization

Frontier of the subject:

(1) Objectifying velocity analysis

(2) Integrating nonlinear physics with velocity analysis

An Optimization Problem for v

Goal: an objective $J[v,d]$ extremized by "correct" v

Well-defined for finite energy $d \Rightarrow J[v,d] = \langle d, W[v]d \rangle + ...$ with $W[v]$ bounded on L^2, $v \in V$.

Problem huge and data noisy $\Rightarrow v, d \mapsto J[v,d]$ differentiable - must be able to use Newton!!!

Theorem (Stolk & WWS, 2003): $v, d \mapsto J[v,d]$ smooth $\Leftrightarrow W[v] = -F[v]R_0[v]\partial_h^2 R_0[v]^* F[v]^*$ with $R_0[v]$ pseudodifferential order -1 ("differential semblance").

Some theory, many successful numerical tests of differential semblance using synthetic and field data: WWS et al., Chauris & Noble 2001, Mulder & tenKroode 2002. Brandsberg-Dahl & De Hoop 2004, Foss et al. 2004.

Ditch Partial Linearization!

Multiply scattered waves frequently evident in field data (phases, apparent velocity) $\Rightarrow$ single scattering (linearization) not an adequate description of seismic wave propagation/reflection.

Two possible resolutions:

(1) process data to remove deviations from single scattering - conventional approach, but ultimately nonphysical

(2) incorporate multiple scattering in inversion, via use of the full nonlinear model.

Current work: extension of differential semblance inversion beyond single scattering via *nonlinear extended models*.

Six Things to Remember

- The seismic inverse scattering problem is a parameter estimation problem.
- Both data and Earth models are broadband, but Earth model contains long scales missing in data. Cross-scale interaction provides both obstacles and opportunities.
- Cross-scale obstacle: connection between long Earth model scales and data h. f. phases $\Rightarrow$ straightforward LS inversion (seismic version of 4DVar) generally unsuccessful.
- Linearized problem about long scale reference model quite tractable, *positions* of short scale components largely independent of noise models.
- Cross-scale opportunity: redundancy of data, extended models $\Rightarrow$ consistency based estimates of long scale model components missing from data.
- Frontier of this subject: reintroduction of fully nonlinear physics into inversion, avoiding pitfalls of naive LS.

Thanks to...

National Science Foundation
Department of Energy
Sponsors of The Rice Inversion Project
Schlumberger for graphics
ExxonMobil and Shell for data
The organizers and IPAM, for inviting me
All of you, for listening

`http://www.trip.caam.rice.edu`