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, pehaps in all directions. More flexible notion needed!!

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

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

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

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

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

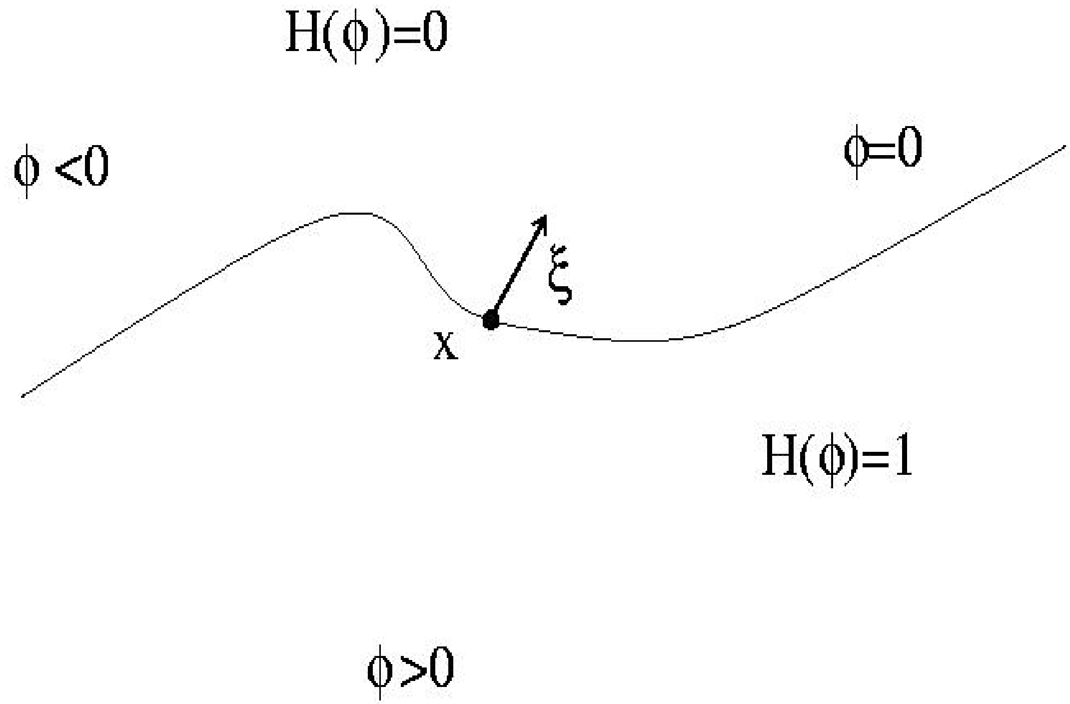
Housekeeping chores:

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

(ii) $WF(u)$ is invariant under chg. of coords 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 $WF(u) \subset \mathcal{N}_f = \{(\mathbf{x}, \xi) : f(\mathbf{x}) = 0, \xi || \nabla f(\mathbf{x})\}$ (*normal bundle* of $f = 0$).



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

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:

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

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

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

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

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

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

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

and start integrating by parts: eventually

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

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

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:

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

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

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

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

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

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

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

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

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!**) $\Rightarrow$

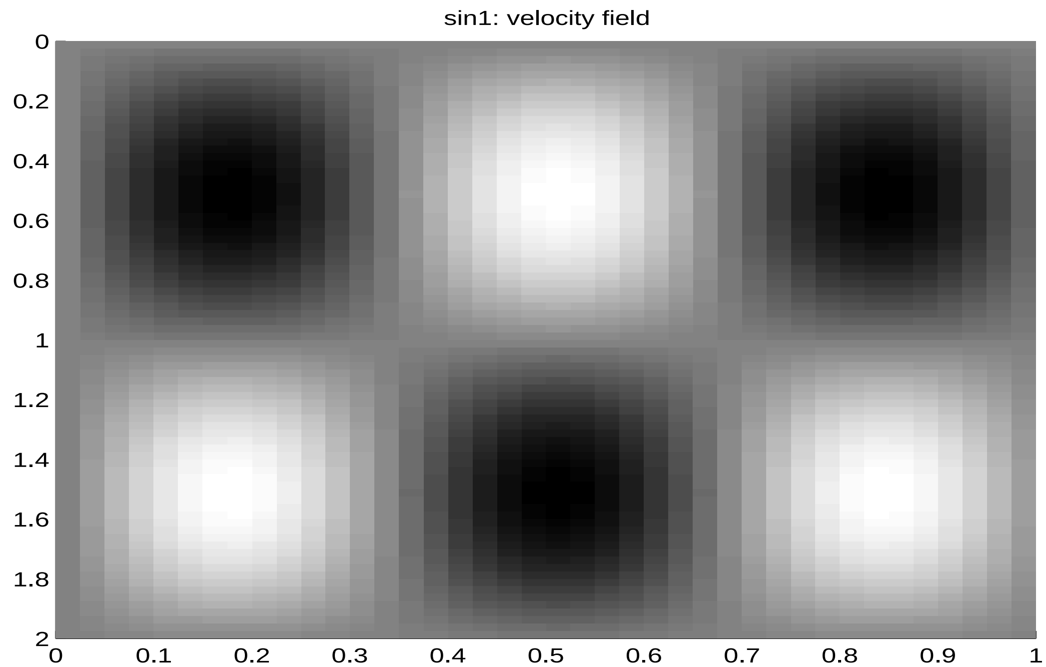
$$\int dx \int_0^T dt\,|R(\mathbf{x},t;\mathbf{x}_s)|^2 \le C\|v\|_{C^4}$$

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.

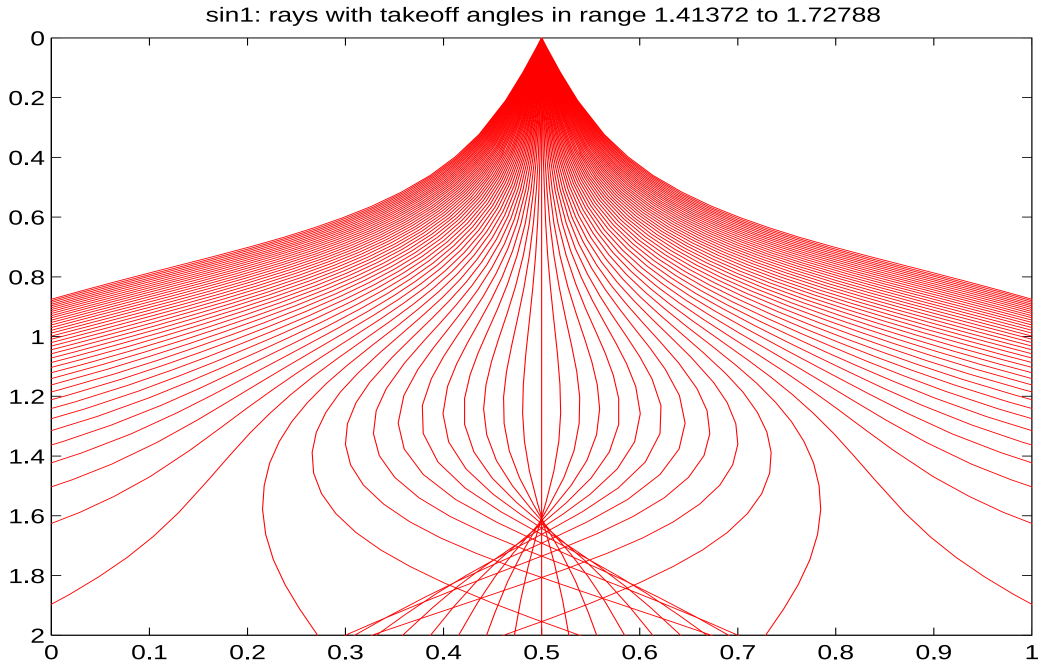
"Not too far" means: there should be one and only one ray of geometric optics connecting each $\mathbf{x}_s$ or $\mathbf{x}_r$ to each $\mathbf{x} \in \mathrm{supp} r$.

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: $\operatorname{supp} r$ 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

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

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

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

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, 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 s-moother, so the first term of the geometric optics approximation (displayed above) suffices (can formalize this with modification of wavefront set defn).

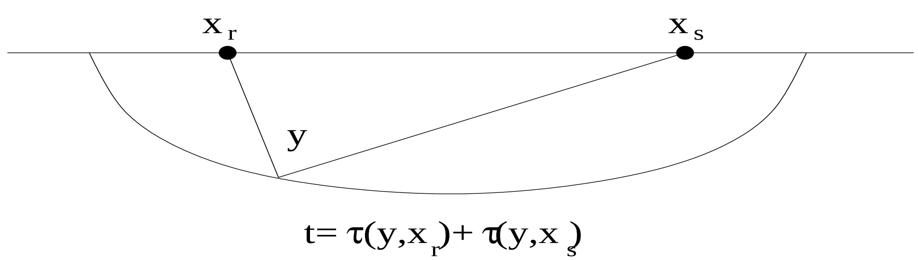
These lectures will ignore the distinction.

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



Zero Offset data and the Exploding Reflector

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

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

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

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

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

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

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

$\sim$ solution w of

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Anatomy of an adjoint:

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

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

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

$N[v]$ = **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**

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

So

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

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

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

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

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

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*.

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* ("ΨDO"):

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

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

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

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

Imaging property of Kirchhoff migration follows from *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. 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 $(\mathbf{x}_0, \xi_0) \notin WF(u)$, 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.

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

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

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

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

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:

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

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

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

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, $|\widehat{\psi u}(\tau\theta)| \le C_N \tau^{-N}$ 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 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

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

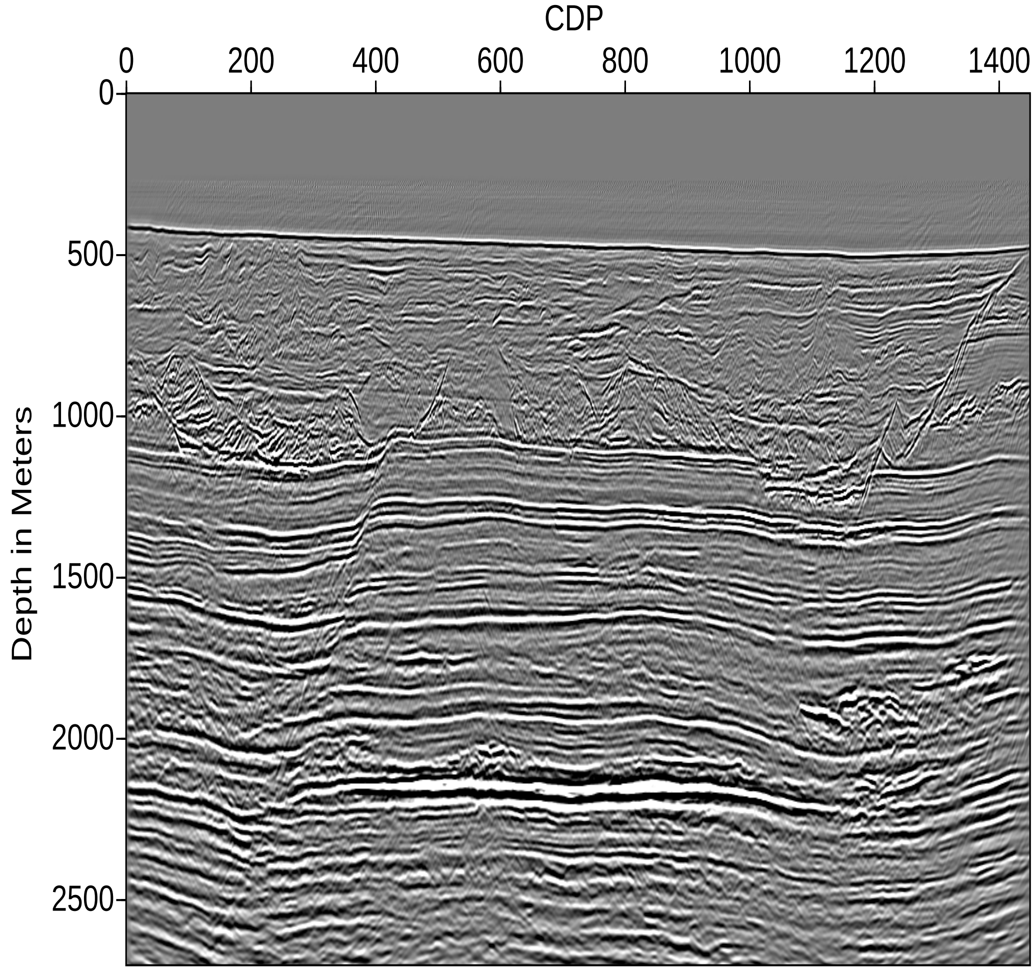
yields $F[v]^\dagger F[v]r \simeq r$ if $WF(r) \subset \Gamma[v]$

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.