2. High frequency asymptotics and imaging operators

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

NB: “everybody’s favorite reflector”: the smooth interface across which $r$ jumps. *But* this is an oversimplification - reflectors in the Earth may be complex zones of rapid change, perhaps in all directions. More flexible notion needed!!
Paley-Wiener characterization of smoothness: \( u \in \mathcal{D}'(\mathbb{R}^n) \) is smooth at \( x_0 \) \iff for some nbhd \( X \) of \( x_0 \), 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, \textit{après} Hörmander: the \textbf{wave front set} \( WF(u) \subset \mathbb{R}^n \times \mathbb{R}^n - 0 \) of \( u \in \mathcal{D}'(\mathbb{R}^n) \) - captures orientation as well as position of singularities.

\((x_0, \xi_0) \notin WF(u) \iff\), there is some open nbhd \( X \times \Xi \subset \mathbb{R}^n \times \mathbb{R}^n - 0 \) of \( (x_0, \xi_0) \) 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 $\Xi$ 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 $T^*(\mathbb{R}^n)$ (i.e. the $\xi$ components transform as covectors).


The standard example: if $u$ jumps across the interface $f(x) = 0$, otherwise smooth, then $WF(u) \subset N_f = \{(x, \xi) : f(x) = 0, \xi \parallel \nabla f(x)\}$ (normal bundle of $f = 0$).
$\phi < 0$

$\phi > 0$

$WF(H(f)) = \{(x, \xi) : f(x) = 0, \xi \parallel \nabla f(x)\}$
Fact ("microlocal property of differential operators"):

Suppose $u \in \mathcal{D}'(\mathbb{R}^n)$, $(x_0, \xi_0) \notin WF(u)$, and $P(x, D)$ is a partial differential operator:

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

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

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

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

Then $(x_0, \xi_0) \notin WF(P(x, D)u)$ [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

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

and start integrating by parts: eventually

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

where $\phi_\alpha \in \mathcal{D}(X)$ is a linear combination of derivatives of $\phi$ and the $a_\alpha$s. Since each integral is rapidly decreasing as $\tau \to \infty$ for $\xi \in \Xi$, it remains rapidly decreasing after multiplication by $\tau^{||\alpha||}$, and so does the sum. Q. E. D.
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
\[
G(x, t; x_s) = \text{retarded fundamental solution:}
\]
\[
\left( 1 \frac{\partial^2}{v^2 \partial t^2} - \nabla^2 \right) G(x, t; x_s) = \delta(t)\delta(x - bx_s)
\]
and \( G \equiv 0, t < 0 \). Then \((w = \delta!\) \)
\[
p = \frac{\partial G}{\partial t}, \quad \delta p = \frac{\partial \delta G}{\partial t}, \quad \text{and}
\]
\[
\left( 1 \frac{\partial^2}{v^2 \partial t^2} - \nabla^2 \right) \delta G(x, t; x_s) = \frac{2}{v^2(x)} \frac{\partial^2 G}{\partial t^2}(x, t; x_s)r(x)
\]
Simplification: from now on, define \( F[v]r = \delta G|_{x=x_r} \) - i.e. lose a \( t \)-derivative. Duhamel’s principle \(\Rightarrow\)
\[
\delta G(x_r, t; x_s) = \int dx \frac{2r(x)}{v(x)^2} \int ds G(x_r, t - s; x) \frac{\partial^2 G}{\partial t^2}(x, s; x_s)
\]
Geometric optics approximation of $G$ should be good, as $v$ is smooth. Local version: if $x$ “not too far” from $x_s$, then

$$G(x, t; x_s) = a(x; x_s)\delta(t - \tau(x; x_s)) + R(x, t; x_s)$$

where the traveltime $\tau(x; x_s)$ solves the eikonal equation

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

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

and the amplitude $a(x; x_s)$ 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!) $\Rightarrow$

$$\int dx \int_0^T dt |R(x, t; x_s)|^2 \leq C\|v\|_C^4$$
Numerical solution of eikonal, transport: ray tracing (Lagrangian), various sorts of upwind finite difference (Eulerian) methods. See Sethian lectures, WWS 1999 MGSS notes (online) for details.

“Not too far” means: there should be one and only one ray of geometric optics connecting each $x_s$ or $x_r$ to each $x \in \text{suppr}$. 

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

Upon caustic formation, the simple geometric optics field description above is no longer correct (Ludwig, 1966).
2D Example of strong refraction: Sinusoidal velocity field \( v(x, z) = 1 + 0.2 \sin \frac{\pi z}{2} \sin 3\pi x \)
Rays in sinusoidal velocity field, source point = origin. Note formation of caustic, multiple rays to source point in lower center.
Assume: $\text{supp } r$ contained in simple geometric optics domain (each point reached by unique ray from any source point $x_s$).

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

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

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

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

provided that $\nabla_x \tau(x, x_r) + \nabla_x \tau(x, x_s) \neq 0 \iff$ velocity at $x$ of ray from $x_s$ not negative of velocity of ray from $x_r \iff$ 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 $C^\infty$ - then the two sides have the same singularities, ie. the same wavefront set.

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

These lectures will ignore the distinction.
So: for \( r \) supported in simple geometric optics domain, no forward scattering \( \Rightarrow \)

\[
\delta G(x_r, t; x_s) \simeq \frac{\partial^2}{\partial t^2} \int dx \frac{2r(x)}{v^2(x)} a(x, x_r) a(x, x_s) \delta(t - \tau(x, x_r) - \tau(x, x_s))
\]

That is: pressure perturbation is sum (integral) of \( r \) over reflection isochron \( \{ x : t = \tau(x, x_r) + \tau(x, x_s) \} \), w. weighting, filtering. Note: if \( v = \text{const.} \) then isochron is ellipsoid, as \( \tau(x_s, x) = |x_s - x|/v! \)
Zero Offset data and the Exploding Reflector

Zero offset data \((x_s = x_r)\) is seldom actually measured (contrast radar, sonar!), but routinely approximated through NMO-stack (to be explained later).

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

Imaging basis: the exploding reflector model (Claerbout, 1970's).
For zero-offset data, distribution kernel of $F[v]$ is

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

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

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

~ solution $w$ of

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

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

18
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

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

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

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

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

(half velocity = double traveltime, same rays!).
Difference between effects of $K, \tilde{G}$: for each $x_s$ 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 \( \delta G \), and the idealized data \( \delta G |_{z=0} \) are really only functions of \( t \) and half-offset \( h = |x_s - x_r|/2 \). 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 \( x_m = (x_r + x_s)/2 \).
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 \( v = v(x_m, z), r = r(x_m, z). \)

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

\[
= \int dz \frac{2r(z)}{v^2(z)} \int d\omega \int dx \omega^2 a(x, x_r) a(x, x_s) e^{i\omega(t - \tau(x, x_r) - \tau(x, x_s))}
\]
Since we have already thrown away smoother (lower frequency) terms, do it again using \textit{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, \( F[v] \) \textit{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:
\[
\int dt \int dh \, d(t, h) F[v] r(t, h) = \int dt \int dh \, d(t, h) A(z(t, h), h) R(z(t, h)) \\
= \int dz \, R(z) \int dh \frac{\partial t}{\partial z}(z, h) A(z, h) d(t(z, h), h) = \int dz \, r(z) (F[v]^* d)(z) \\
\text{so } F[v]^* = -\frac{\partial}{\partial z} S M[v] N[v],
\]

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

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

\[S = \textbf{stacking operator } \quad S f(z) = \int dh \, f(z, h)\]
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 \( x_m \) and you’ve got standard processing. (end of layered interlude).
**Multioffset Imaging:** if \( d = F[v]r \), then
\[
\]
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”):

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

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

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

Explicit computation of symbol $p$ - for details, .
Imaging property of Kirchhoff migration follows from microlocal property of $\Psi DOs$:

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

Will prove this. First, a few other properties:

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

Complete accounts of theory, many apps: books of Duistermaat, Taylor, Nirenberg, Treves, Hörmander.
Proof of microlocal property: suppose \((x_0, \xi_0) \notin WF(u)\), choose neighborhoods \(X, \Xi\) as in defn, with \(\Xi\) conic. Need to choose analogous nbhds for \(P(x, D)u\). Pick \(\delta > 0\) so that \(B_{3\delta}(x_0) \subset X\), set \(X' = B_\delta(x_0)\).

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 \(\mathcal{F}(\phi P(x, D)u)\). Choose \(\psi \in \mathcal{E}(X)\) so that \(\psi \equiv 1\) on \(B_{2\delta}(x_0)\).

NB: this implies that if \(x \in X', \psi(y) \neq 1\) then \(|x - y| \geq \delta\).
Write \( u = (1 - \psi)u + \psi u \). Claim: \( \phi P(x, D)((1 - \psi)u) \) is smooth.

\[
\phi(x)P(x, D)((1 - \psi)u))(x) = \phi(x) \int d\xi P(x, \xi)e^{ix\cdot\xi} \int dy (1 - \psi(y))u(y)e^{-iy\cdot\xi}
\]

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

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

using the identity

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

and integrating by parts \(2M\) times in \(\xi\). This is permissible because \(\phi(x)(1 - \psi(y)) \neq 0 \Rightarrow |x-y| > \delta\).
According to the definition of $\Psi DO$,

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

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 $\phi P(x, D)(\psi u)$. Pick $\eta \in \Xi'$ and w.l.o.g. scale $|\eta| = 1$. Fourier transform:

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

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

$$= \tau^n \int dx \int d\theta P(x, \tau \theta) \phi(x) \hat{\psi}u(\tau \theta) e^{i\tau x \cdot (\theta - \eta)}$$
Divide the domain of the inner integral into \( \{ \theta : |\theta - \eta| > \epsilon \} \) and its complement. Use

\[
-\nabla^2_x e^{i\tau x \cdot (\theta - \eta)} = \tau^2 |\theta - \eta|^2 e^{i\tau x \cdot (\theta - \eta)}
\]

and integration by parts \( 2M \) times to estimate the first integral:

\[
\tau^{n-2M} \left| \int dx \int_{|\theta - \eta| > \epsilon} d\theta \left(-\nabla^2_x\right)^M [P(x, \tau \theta) \phi(x)] \tilde{\psi} u(\tau \theta) \right. \\
\times \left. |\theta - \eta|^{-2M} e^{i\tau x \cdot (\theta - \eta)} \right| \\
\leq C \tau^{n+m-2M}
\]

\( m \) being the order of \( P \). Thus the first integral is rapidly decreasing in \( \tau \).
For the second integral, note that $|\theta - \eta| \leq \epsilon \Rightarrow \theta \in \Xi$, per the defn of $\Xi'$. Since $X \times \Xi$ is disjoint from the wavefront set of $u$, for a sequence of constants $C_N$, $|\hat{\psi}u(\tau\theta)| \leq C_N\tau^{-N}$ uniformly for $\theta$ in the (compact) domain of integration, whence the second integral is also rapidly decreasing in $\tau$. \textbf{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.

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** $\Gamma[v] \subset \mathbb{R}^3 \times \mathbb{R}^3 - 0$: 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 $b(x_r, t; x_s)$, the modified Kirchhoff migration operator

$$F[v]^\dagger d(x) = \int \int \int dx_r dx_s dt b(x_r, t; x_s) \delta(t - \tau(x; x_s) - \tau(x; x_r)) d(x_r, t; x_s)$$

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.