Imaging in Random Media

George Papanicolaou
Department of Mathematics, Stanford University
http://georgep.stanford.edu

Outline

- Monday I. Introduction and overview
  II. Time reversal, migration and least squares imaging. Basic resolution theory
  Monday afternoon: C. Tsogka. An overview of computational results with migration imaging and introduction to computational issues

- Tuesday III. Resolution theory, use of the Kirchhoff-Helmholtz identities
  IV. Noise sources and correlations. Open media and cavities. Velocity estimation and imaging with distributed sensors (to be continued on Thursday afternoon by J. Garnier).
  Tuesday afternoon: C. Tsogka. Computational wave propagation and array imaging

Outline continued

- Wednesday V. The singular value decomposition, in detection and imaging
  VI. Edge illumination, the Fraunhofer regime and inverse filters
  Wednesday afternoon: L. Borcea. Imaging with layer annihilation

- Thursday VII. Waves in random media: Layered media, the paraxial approximation, radiative transport
  VIII. Time reversal in random media, super-resolution, statistical stability
  Thursday afternoon: J. Garnier. Passive sensor imaging with cross correlations

- Friday IX. Coherent interferometry for imaging in random media
  X. Discussion of research problems: Time reversal, imaging, random media, simulations, communications, optimization and adaptivity

Introduction and overview

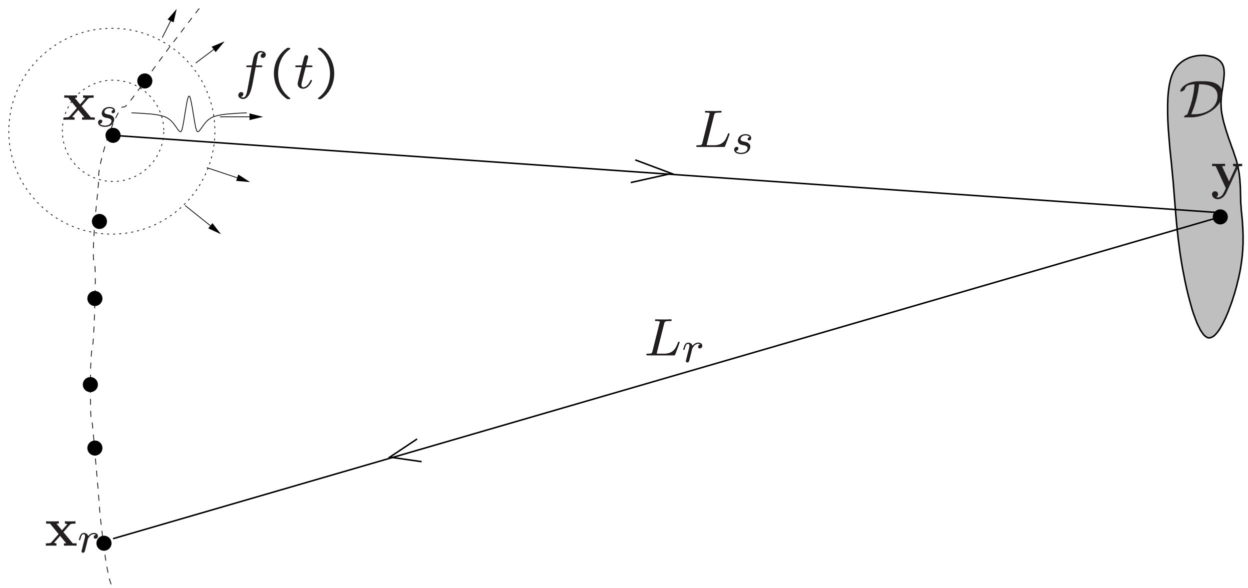
Inverse problems, detection, imaging

- Inverse problem: Given data as part of the solution of a problem (PDE ...), find the unknown parameters (structure) in it. Overly general, too inclusive for imaging

- Detection: Given two (or more) sets of data and an underlying model, find if they are consistent with it. Notion of detectability threshold. Too special and restrictive for imaging

- Imaging: From imperfect information (rough forward models, limited and noisy data) estimate parts of the unknown (parameter) structure that is of interest. In particular, quantify and understand the trade-offs between data size, computational complexity, and resolution.

Passive and Active Sensor data



Active sensor data: $P(\mathbf{x}_r, \mathbf{x}_s, t)$ for $(\mathbf{x}_r, \mathbf{x}_s, t)$ a set of source-receiver locations. Can be up to a function in $R^2 \times R^2$ plus time in R_+ for planar arrays.

Passive sensor data: $P(\mathbf{x}_r, t)$. can be up to a three-dimensional dataset for planar arrays.

Acquisition geometries, data structure

Different data acquisition geometries: Arrays, synthetic aperture arrays (zero-offset), distributed sensors, full aperture imaging (medical).

Narrowband (microwaves), broadband (ultrasound) and noise probing signals.

Coherent (radar, sonar, seismic ...) and incoherent (infrared, optical tomography, X-ray tomography) imaging.

Signal-to-Noise ratio (SNR) issues.

The time Fourier transform of the data, $\widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega)$, is modeled by $\widehat{U} = \widehat{f}_B(\omega - \omega_0)\widehat{G}_F(\mathbf{x}_r, \mathbf{x}_s, \omega; c)$ where $\widehat{f}_B(\omega - \omega_0)$ is the Fourier transform of the pulse $f(t) = e^{-i\omega_0 t} f_B(t)$ and $\widehat{G}_F$ is the Green's function that solves the Helmholtz equation

$$(\Delta + k^2 n^2(\mathbf{x}))\widehat{G}_F = -\delta(\mathbf{x} - \mathbf{y}), \quad k = \frac{\omega}{c_0}, \quad n(\mathbf{x}) = \frac{c_0}{c(\mathbf{x})}$$

with a radiation condition. The index of refraction is $n(\mathbf{x})$.

The inverse problem, the array least squares problem, is:
Minimize $J[c] + \alpha ||c||_{REG}$ where

$$J[c] = \int d\omega \sum_{\mathbf{x}_s, \mathbf{x}_r} \left| \widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega) - \widehat{U}(\mathbf{x}_r, \mathbf{x}_s, \omega; c) \right|^2$$

and α is a strength of regularization parameter.

Note that this is a NONLINEAR problem for the unknown index of refraction $n(x)$ or propagation speed $c(x)$.

Structure of velocity (refractive index) and linearization

We need a model that distinguishes between
(a) a background velocity that is known or can be estimated,
(b) the reflectors or targets that we wish to image, and
(c) the clutter that is part of the background that we do not know, and can only estimate its overall influence statistically.

This motivates writing

$$n^2(\mathbf{x}) = n_{BG}^2(\mathbf{x}) + \rho(\mathbf{x}) + \mu(\mathbf{x})$$

where the background index of refraction is $n(\mathbf{x}) = c_0/c_{BG}(\mathbf{x})$, the target reflectivity is $\rho(\mathbf{x})$, and the clutter is modeled by the mean zero, stationary random function $\mu(\mathbf{x})$.

We next linearize in the reflectivity by writing $\widehat{G}_F = \widehat{G} + \delta\widehat{G}$ where

$$(\Delta + k^2(n_{BG}^2(\mathbf{x}) + \mu(\mathbf{x})))\widehat{G} = -\delta(\mathbf{x} - \mathbf{y})$$

is the (random) background Green's function and

$$\delta\widehat{G}(\mathbf{x}, \mathbf{y}, \omega) = k^2 \int d\mathbf{z} \rho(\mathbf{z}) \widehat{G}(\mathbf{x}, \mathbf{z}, \omega) \widehat{G}(\mathbf{z}, \mathbf{y}, \omega)$$

We model the data by

$$\widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega) \sim \widehat{f}_B(\mathbf{x}_s, \omega - \omega_0)\delta\widehat{G}(\mathbf{x}_r, \mathbf{x}_s, \omega) = (\widehat{A}(\omega)\rho)(\mathbf{x}_r, \mathbf{x}_s)$$

where $\widehat{A}(\omega)$ is the random, frequency dependent, linear operator that maps reflectivities to array data. The least squares linearized inverse problem is to minimize $J_L[\rho]$ where

$$J_L[\rho] = \int d\omega \sum_{\mathbf{x}_r, \mathbf{x}_s} |\widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega) - (\widehat{A}(\omega)\rho)(\mathbf{x}_r, \mathbf{x}_s)|^2$$

When ρ is a sum of M functions with small support (compared to the wavelength) with M smaller than the array size N, then the normal solution for this problem is

$$\rho \sim \int d\omega (\widehat{A}^H(\omega)\widehat{A}(\omega))^{-1}\widehat{A}^H(\omega)\widehat{P}(\omega)$$

In the general case

$$\rho \sim \int d\omega \widehat{A}^H(\omega)(\widehat{A}(\omega)\widehat{A}^H(\omega))^{-1}\widehat{P}(\omega)$$

The adjoint operator $\widehat{A}^H(\omega)$ maps array data to reflectivities and is given by

$$(\widehat{A}^H(\omega)\widehat{P}(\omega))(\mathbf{z}) = \frac{\omega^2}{c_0^2}\sum_{\mathbf{x}_s,\mathbf{x}_r} \widehat{f}_B(\mathbf{x}_s,\omega-\omega_0)\widehat{G}(\mathbf{z},\mathbf{x}_s,\omega)\widehat{G}(\mathbf{z},\mathbf{x}_r,\omega) \cdot\overline{\widehat{P}(\mathbf{x}_r,\mathbf{x}_s,\omega)}$$

Problem: $\widehat{A}^H(\omega)$ is random and is not known, so this least squares solution cannot be implemented to give an image!
However: The operator $\int d\omega \widehat{A}^H(\omega)\widehat{A}(\omega)$ is close to the identity operator, with high probability. It is related to the time reversal operator for source imaging. A similar result holds for $\int d\omega \widehat{A}(\omega)\widehat{A}^H(\omega)$. This motivates dropping the normalizing factors (the inverses) in the least squares solution

$$\rho(\mathbf{z}) \sim \int d\omega(\widehat{A}^H(\omega)\widehat{P}(\omega))(\mathbf{z}),$$

but resolution is affected, especially in narrowband cases and in random media!

Kirchhoff or travel time migration

Assume that the background velocity is known and that there is no clutter. Denote the deterministic background Green's function by $\widehat{G}_0(\mathbf{x}, \mathbf{y}, \omega) = \frac{e^{i\omega\tau(\mathbf{x},\mathbf{y})}}{4\pi|\mathbf{x}-\mathbf{y}|}$. We can then use the following imaging functional for the reflectivity $\rho(\mathbf{y}^S)$:

$$I^{KM}(\mathbf{y}^S) = \sum_{\mathbf{x}_s, \mathbf{x}_r} P(\mathbf{x}_r, \mathbf{x}_s, \tau(\mathbf{x}_s, \mathbf{y}^S) + \tau(\mathbf{y}^S, \mathbf{x}_r))$$

Here $\tau(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|/c_0$ is the travel time from $\mathbf{x}$ to $\mathbf{y}$ when the speed of propagation is c_0.

Travel time migration (1970's) is an elegant way to 'triangulate' the location of a scatterer using array (or distributed sensor) data without having to estimate travel times.

Mathematical theory for travel time migration has been developed by Beylkin, Burridge, Symes, Bleistein, ...

Denote the deterministic background Green's function by $\widehat{G}_0(\mathbf{x}, \mathbf{y}, \omega)$. Then the migration functional for imaging the reflectivity $\rho(\mathbf{y}^S)$ can be written as

$$I(\mathbf{y}^S) = \int d\omega \sum_{\mathbf{x}_s, \mathbf{x}_r} \overline{\widehat{P}(\mathbf{x}_r, \mathbf{x}_s, \omega)} \widehat{G}_0(\mathbf{y}^S, \mathbf{x}_s, \omega) \widehat{G}_0(\mathbf{y}^S, \mathbf{x}_r, \omega)$$

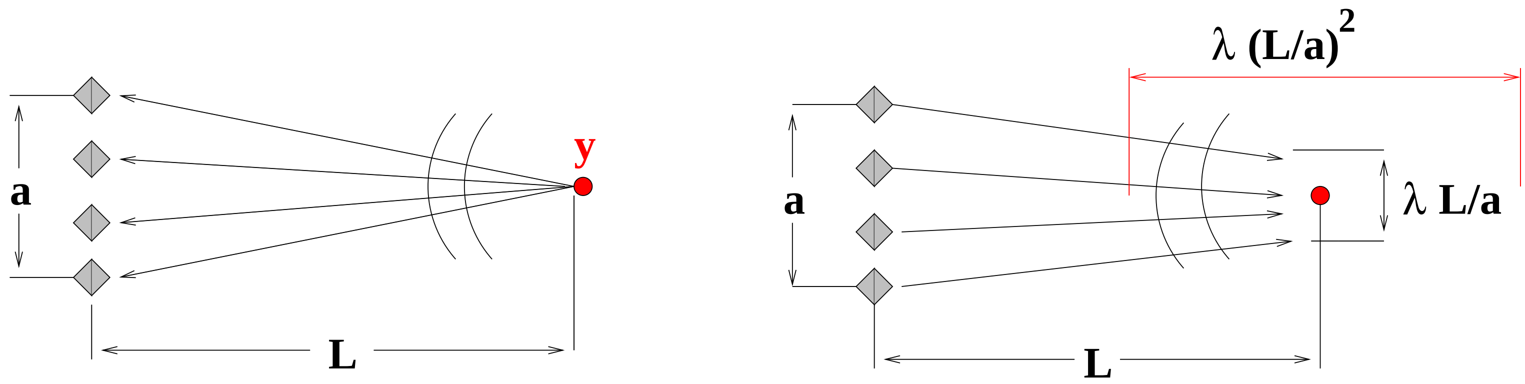
If we take $\widehat{G}_0 \sim e^{i\omega\tau(\mathbf{x},\mathbf{y})}$, where $\tau(\mathbf{x}, \mathbf{y})$ is the travel time from $\mathbf{x}$ to $\mathbf{y}$, then in the time domain we get the Kirchhoff migration imaging functional.

For active sources (passive arrays) the imaging functional is

$$I(\mathbf{y}^S) = \int d\omega \sum_{\mathbf{x}_r} \overline{\widehat{P}(\mathbf{x}_r, \omega)} \widehat{G}_0(\mathbf{y}^S, \mathbf{x}_r, \omega),$$

which is the time reversed field in homogeneous (unphysical) medium.

Basic resolution theory in homogeneous media.



Cross-range (Rayleigh) resolution: $\frac{\lambda_0 L}{a}$

Range resolution (narrowband): $\frac{\lambda_0 L^2}{a^2}$

Range resolution (broadband): $\frac{c_0}{B}$

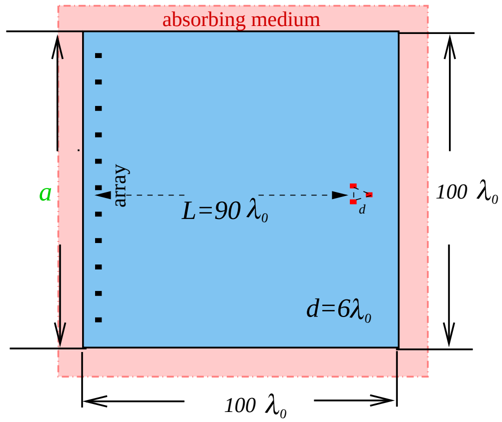
Ultrasonic nondestructive testing: $\lambda_0 = 3cm,\ a = 1m,\ L = 3 - 5m$ (propagation speed $3Km/sec.$)

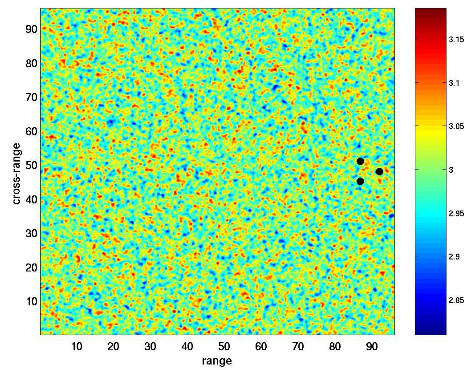
Cross-range resolution: $9 - 15cm$

- What happens in a randomly inhomogeneous medium?

It depends on whether it is known exactly, as in TIME-REVERSAL, or only its large scale features are known, as in IMAGING.

Numerical simulations for imaging in opaque structures

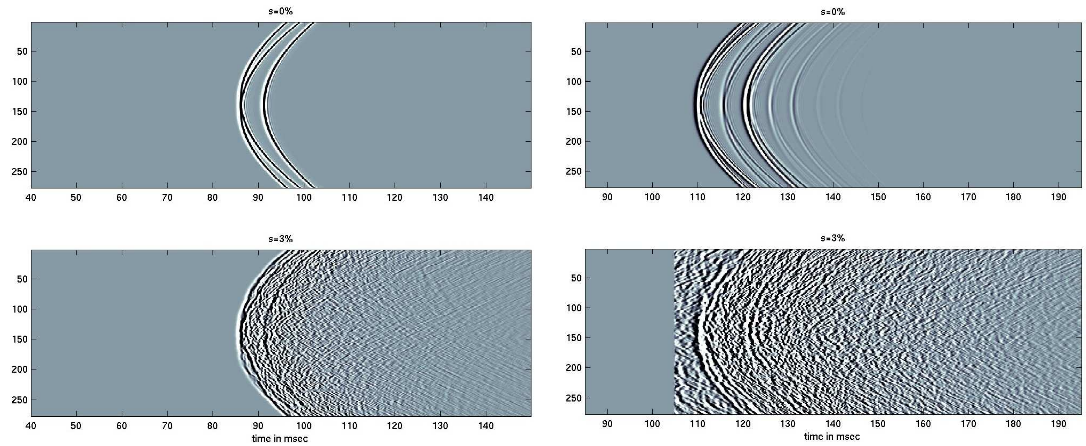




Computational domain $100\lambda_0 \times 100\lambda_0$ with central wavelength $\lambda_0 = 3cm$ (at central frequency $f_0 = 100KHz$ and with $c_0 = 3km/sec$), surrounded by a perfectly matched layer (pink).

The array has 185 receiving elements $\lambda_0/2$ apart, for an aperture of $92\lambda_0$.

-Passive array -Active array



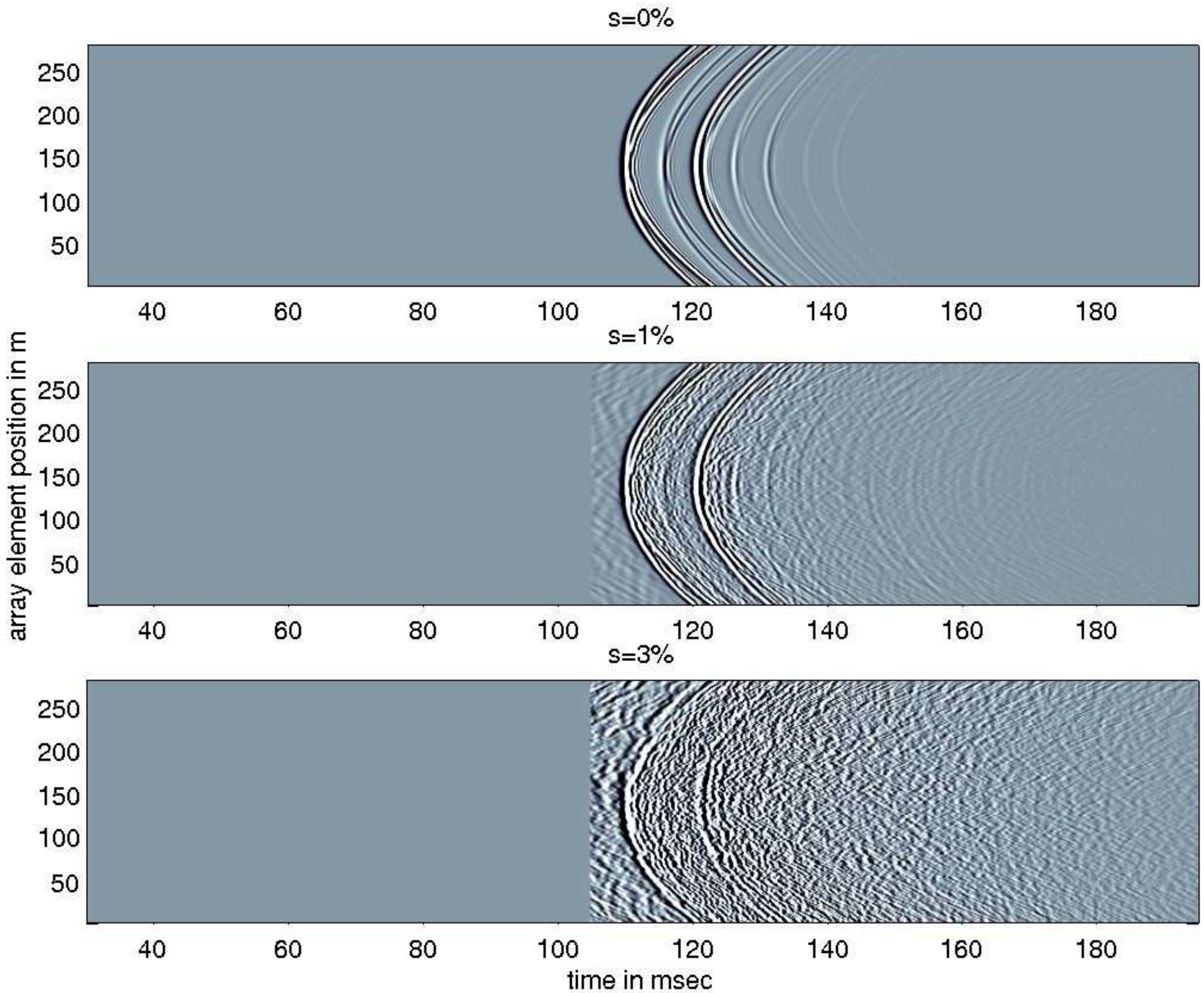
Left figures: Passive array

Right figures: Active array with central illumination

Top figures: Homogeneous medium.

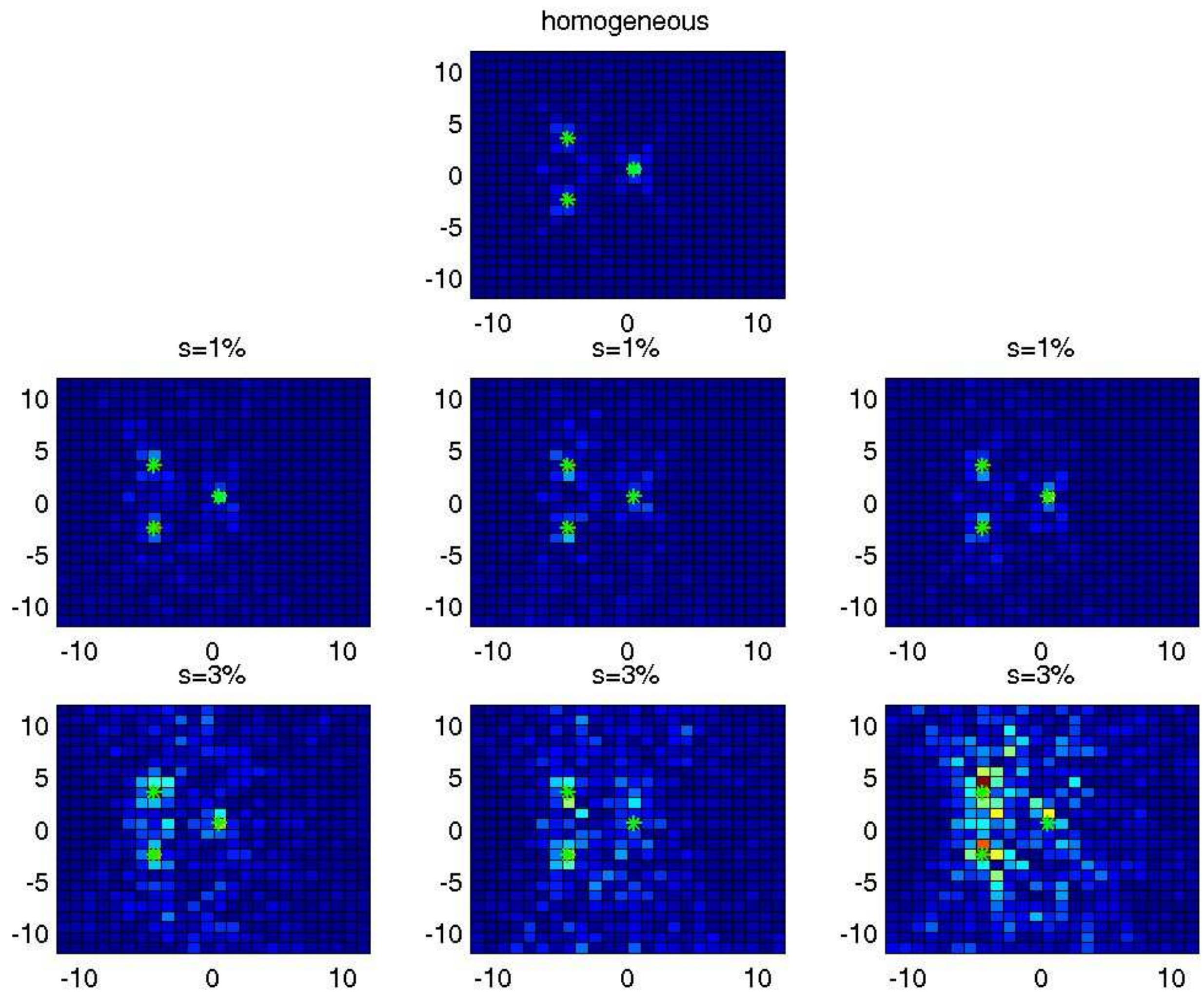
Bottom figures: Random medium with with standard deviation $s = 3\%$.

Why only main (Born) scattering matters in clutter



Down: Homogeneous, 1%, 3%STD.

Kirchhoff migration or travel time imaging (passive array)



Down: Homogeneous, 1%, 2%, 3%STD. Across: different realizations.

Critique of travel time migration imaging

It does not work well in clutter.

It is statistically unstable in clutter.

The reason is that KM tries to cancel the random phase of the signals arriving at the array with a deterministic phase using travel times.

The true Green's function for the random medium is not known and so cannot be used for imaging, which would result is huge resolution enhancement as in physical time reversal.

Time reversal, migration and least squares imaging:

Basic resolution theory

Point spread function in a homogeneous medium

Recall the Fourier transform:

$$\hat{g}(\omega) = \int_R e^{i\omega t} g(t) dt \ ; \quad g(t) = \frac{1}{2\pi} \int e^{-i\omega t} \hat{g}(\omega) d\omega$$

The FT of the pulse $e^{-i\omega t} f_B(t)$ (real part) emitted by the source is $\hat{f}_B(\omega - \omega_0)$ with $\hat{f}_B = 0$ for $\omega > B/2$. The bandwidth is B.

A point source at $\mathbf{y}$ gives the array data

$$P(\mathbf{x}_r, t) = e^{-i\omega_0 t} \frac{f_B(t - \tau(\mathbf{x}_r, \mathbf{y}))}{4\pi |\mathbf{x}_r - \mathbf{y}|} \ , \ \hat{P}(\mathbf{x}_r, \omega) = \hat{f}_B(\omega - \omega_0) \frac{e^{i\omega\tau(\mathbf{x}_r, \mathbf{y})}}{4\pi |\mathbf{x}_r - \mathbf{y}|}$$

where $\tau(\mathbf{x}_r, \mathbf{y})$ is the travel time from $\mathbf{y}$ to $\mathbf{x}_r$.

For a very broadband pulse (impulse response) the signal arrives at $\mathbf{x}_r$ at this travel time. Therefore $|\mathbf{x}_r - \mathbf{y}|^2 = c_0^2 t^2$. If L is the distance of $\mathbf{y}$ to the array and x_r is the distance from $\mathbf{x}_r$ to the nearest point from $\mathbf{y}$ to the array, then $x_r^2 + L^2 = c_0^2 t^2$ or

$$\frac{t^2}{L^2/c_0^2} - \frac{x_r^2}{L^2} = 1$$

which is a hyperbola in (x_r, t).

Place a point source ar $\mathbf{x}_r$ and let $g(t, \mathbf{x}_r)$ be the pulse emitted from it. The signal received at a search point $\mathbf{y}^S$ is

$$\sum_{\mathbf{x}_r} \frac{g(t - \tau(\mathbf{x}_r, \mathbf{y}), \mathbf{x}_r)}{4\pi|\mathbf{x}_r - \mathbf{y}^S|}$$

How do we choose $g(t, \mathbf{x}_r)$ so as to beamform a pulse to $\mathbf{y}$?

Use time reversal: $g(t, \mathbf{x}_r) = P(\mathbf{x}_r, -t)$. Then the signal at $\mathbf{y}^S$ is

$$e^{-i\omega_0 t} \sum_{\mathbf{x}_r} \frac{f_B(-t + \tau(\mathbf{x}_r, \mathbf{y}) - \tau(\mathbf{x}_r, \mathbf{y}^S))}{(4\pi)^2|\mathbf{x}_r - \mathbf{y}^S|^2}$$

and in the Fourier domain

$$\overline{\widehat{f}_B(\omega - \omega_0)} \sum_{\mathbf{x}_r} \overline{\widehat{G}_0(\mathbf{x}_r, \mathbf{y}, \omega)} \widehat{G}_0(\mathbf{x}_r, \mathbf{y}^S, \omega)$$

How well does this focus around $\mathbf{y}$?

The travel time migration functional is

$$I^{KM}(\mathbf{y}^S) = \int d\omega \sum_{\mathbf{x}_r} \widehat{P}(\mathbf{x}_r,\omega)e^{-i\omega\tau(\mathbf{x}_r,\mathbf{y}^S)} = \sum_{\mathbf{x}_r} P(\mathbf{x}_r,\tau(\mathbf{x}_r,\mathbf{y}^S))$$

and up to multiplicative factors this is the conjugate of beamforming.

Physical time reversal is not an imaging functional but a physical process:

$$\Gamma^{TR}(\mathbf{y}^S) = \int d\omega \sum_{\mathbf{x}_r} \overline{\widehat{P}(\mathbf{x}_r,\omega)}\widehat{G}(\mathbf{x}_r,\mathbf{y}^S,\omega)$$

There is no difference between them in a homogeneous medium.

Basic fact:
I^{KM} loses resolution in random media
Γ^{TR} gains resolution in random media!

Time harmonic point spread function

Detailed analysis of

$$I = \sum_{\mathbf{x}_r} e^{ik(|\mathbf{x}_r - \mathbf{y}| - |\mathbf{x}_r - \mathbf{y}^S|)} \ , \quad k = \frac{\omega}{c_0} \ , \quad \lambda = \frac{2\pi}{k}$$

under the following conditions. If the origin of coordinates is at the center of a linear array, $\mathbf{y} = (L, 0)$, $\mathbf{y}^S = (L + \eta, \xi)$, and $\mathbf{x}_r = (0, rh/2)$ for $r = 0, \pm 1, \pm 2, ..., \pm N$, let $a = Nh$ and assume that $a \ll L$. Assume also that $\lambda \ll a$ and that the spacing $h/2$ between sensors smaller than a half wavelength, $h < \lambda$ so that $\mathbf{x}_r = (0, rh/2) = (0, x)$ with $-a/2 \le x \le a/2$.

We then have:

$$|\mathbf{x}_r - \mathbf{y}| = (L^2 + x^2)^{1/2} = L(1 + (\frac{x}{L})^2)^{1/2} \approx L + \frac{x^2}{2L}$$

and similarly

$$|\mathbf{x}_r - \mathbf{y}^S| = ((L+\eta)^2 + (x - \xi)^2)^{1/2} \approx L + \eta + \frac{(x-\xi)^2}{2(L+\eta)}$$

so that

$$|\mathbf{x}_r - \mathbf{y}| - |\mathbf{x}_r - \mathbf{y}^S| \approx -\eta - \frac{\xi^2}{2(L+\eta)} + \frac{x\xi}{L+\eta} + \frac{\eta\xi^2}{2L(L+\eta)}$$

The sum above can be approximated by an integral

$$I \approx \frac{2}{h} e^{-ik(\eta + \frac{\xi^2}{2(L+\eta)})} \int_{-a/2}^{a/2} e^{ik(\frac{x\xi}{L+\eta} + \frac{\eta\xi^2}{2L(L+\eta)})} dx$$

Continue analysis of time harmonic psf

Change variables in the integral, $x = ax'$. After dropping primes and taking absolute values the integral is

$$\frac{2a}{h}\left|\int_{-1/2}^{1/2} e^{i\pi\left(\frac{ka\xi}{\pi(L+\eta)}x+\frac{k}{2\pi}\frac{a^2\eta}{2L(L+\eta)}x^2\right)}\ dx\right|$$

But $\lambda = 2\pi/k$ so the exponent is

$$\frac{2\xi a}{\lambda(L+\eta)}x + \frac{\eta a^2}{\lambda L(L+\eta)}x^2$$

For $\eta \ll L$ this simplifies to

$$2\frac{\xi}{\lambda L/a}x + \frac{\eta}{\lambda(L/a)^2}x^2$$

We see that the cross range coordinate ξ scales with $\lambda L/a$ and the range resolution with η with $\lambda(L/a)^2$. These are the classical time-harmonic resolution limits in array (aperture) imaging, in the regime $\lambda \ll a \ll L$.

It is simply c_0/B where B is the bandwidth.

This can be seen from the formula

$$e^{-i\omega_0 t} \sum_{\mathbf{x}_r} \frac{f_B(-t + \tau(\mathbf{x}_r, \mathbf{y}) - \tau(\mathbf{x}_r, \mathbf{y}^S))}{(4\pi)^2 |\mathbf{x}_r - \mathbf{y}^S|^2}$$

after noting the width of f_B is proportional to that of its Fourier transform, which is B.

In units of length it is c_0/B.

Full aperture resolution (time harmonic)

What is the resolution of the Kirchhoff migration (or time reversal) functional in a homogeneous medium when the array encloses the source point? Lets consider time reversal

$$\Gamma^{TR}(\mathbf{y}^S,\omega) = \sum_{\mathbf{x}_r \in \partial D} \widehat{P}(\mathbf{x}_r,\omega)\overline{\widehat{G}_0(\mathbf{x}_r,\mathbf{y}^S,\omega)} \approx \int_{\partial D} dS(x)\widehat{P}(\mathbf{x},\omega)\overline{\widehat{G}_0(\mathbf{x},\mathbf{y}^S,\omega)}$$

Assume that D is a convex region and that

$$\frac{|\mathbf{y}-\mathbf{y}^S|}{|\mathbf{x}-\mathbf{y}^S|} \ll 1$$

This condition says that the search point and the source are away from the array, which is the boundary of D. We will see how to generalize this in lecture III.

We will show that the resolution is $\lambda/2$, which is a well known result.

We can get an approximate expression for the difference of the distance to the array from $\mathbf{y}$ and from $\mathbf{y}^S$ under the assumption that they are relatively far from it. We first write

$$\Gamma^{TR}(\mathbf{y}^S,\omega) = \int_{\partial D} dS(x) \frac{e^{ik(|\mathbf{x}-\mathbf{y}|-|\mathbf{x}-\mathbf{y}^S|)}}{(4\pi)^2|\mathbf{x}-\mathbf{y}||\mathbf{x}-\mathbf{y}^S|}$$

Then we note that

$$|\mathbf{x}-\mathbf{y}| - |\mathbf{x}-\mathbf{y}^S| \approx (\mathbf{y}-\mathbf{y}^S)\cdot(\frac{\mathbf{y}^S-\mathbf{x}}{|\mathbf{y}^S-\mathbf{x}|} + o(1))$$

which means that we have to evaluate the integral

$$\Gamma^{TR}(\mathbf{y}^S,\omega) \approx \int_{\partial D} dS(x) \frac{e^{ik(\mathbf{y}-\mathbf{y}^S)\cdot\frac{\mathbf{y}^S-\mathbf{x}}{|\mathbf{y}^S-\mathbf{x}|}}}{(4\pi)^2|\mathbf{x}-\mathbf{y}^S|^2}$$

where we have simplified the denominator as well.

Full aperture psf, continued

Since we have assumed that array ∂D is convex, we can parametrize it with polar coordinates relative to a fixed point x^* on it:

$$|\mathbf{x} - \mathbf{y}^S| = g(\theta, \phi)$$

with θ and ϕ the polar and azimuthal angles, respectively. In these coordinates the surface element has the form $dS(x) = |\mathbf{x} - \mathbf{y}^S|^2 \sin\theta d\theta d\phi$. Therefore

$$\Gamma^{TR}(\mathbf{y}^S, \omega) \approx \int_0^\pi \int_0^{2\pi} \frac{e^{ik|\mathbf{y}^S - \mathbf{y}| \cos\theta}}{(4\pi)^2} \sin\theta d\theta d\phi$$

The integration gives

$$\Gamma^{TR}(\mathbf{y}^S, \omega) \approx \frac{\sin(k|\mathbf{y}^S - \mathbf{y}|)}{4\pi|\mathbf{y}^S - \mathbf{y}|}.$$

We get a resolution estimate from the first zero of the sinc function, the Rayleigh resolution. We have $k|\mathbf{y}^S - \mathbf{y}| = \pi$ or $|\mathbf{y}^S - \mathbf{y}| = \lambda/2$, which is a well known result.

So far we have considered only time harmonic point spread functions. We will now consider a planar array $A \subset R^2$ and signals with bandwidth B, and we will analyze the weighted Kirchhoff migration functional in the limit $A \to R^2$, $B \to \infty$. The imaging functional is

$$I^{WKM}(\mathbf{y}^S; A, B) = \int_{R^3} \rho(\mathbf{z}) \int_A \int_{|\omega-\omega_0|\leq B/2} M(\mathbf{x}, \mathbf{y}^S, \omega)$$

$$\times \frac{e^{i(\omega/c_0)(|\mathbf{x}-\mathbf{z}|-|\mathbf{x}-\mathbf{y}^S|)}}{(4\pi)^2|\mathbf{x}-\mathbf{z}||\mathbf{x}-\mathbf{y}^S|} d\mathbf{z} d\mathbf{x} d\omega$$

The form of the multiplier M is given below. It compensates for the fact that the array is large. It significance will become clearer in Lecture III.

We assume, as in the full aperture, time harmonic case, that

$$\frac{|\mathbf{y}-\mathbf{y}^S|}{|\mathbf{x}-\mathbf{y}^S|} \ll 1$$

We use this condition it to simplify the imaging functional

$$I^{WKM}(\mathbf{y}^S; A, B) \approx \int_{R^3} \rho(\mathbf{z}) \int_A \int_{|\omega-\omega_0|\le B/2} M(\mathbf{x}, \mathbf{y}^S, \omega) \frac{e^{i(\omega/c_0)(\mathbf{z}-\mathbf{y}^S)\cdot\frac{\mathbf{x}-\mathbf{y}^S}{|\mathbf{x}-\mathbf{y}^S|}}}{(4\pi)^2|\mathbf{x}-\mathbf{y}^S|^2}$$

We introduce the change of variables from R^3 to itself and its Jacobian

$$(\mathbf{x}, \omega) \in R^3 \to \zeta = \frac{\omega}{c_0}\frac{\mathbf{x}-\mathbf{y}^S}{|\mathbf{x}-\mathbf{y}^S|} \,, \quad d\mathbf{x}d\omega = \frac{\partial(\mathbf{x},\omega)}{\partial\zeta} = J(\mathbf{x}, \mathbf{y}^S, \omega)d\zeta$$

which is one-to-one and onto as $A \to R^2$ and $B \to \infty$. Let $\mathcal{Z}(A, B)$ be the image of the $(\mathbf{x}, \omega)$ region of integration in ζ space.

With this change of variables we have

$$I^{WKM}(\mathbf{y}^S; A, B) \approx \int_{R^3} \rho(\mathbf{z}) \int_{\mathcal{Z}(A,B)} \frac{e^{i(\mathbf{z}-\mathbf{y}^S)\cdot\zeta}}{(2\pi)^3}$$

provided we choose the multiplier M so that

$$\frac{M(\mathbf{x}, \mathbf{y}^S, \omega) J(\mathbf{x}, \mathbf{y}^S, \omega)}{(4\pi)^2 |\mathbf{x} - \mathbf{y}^S|^2} = \frac{1}{(2\pi)^3}$$

Now as $A \to R^2$ and $B \to \infty$ the inner integral becomes a 3D delta function and therefore

$$I^{WKM}(\mathbf{y}^S; A, B) \approx \rho(\mathbf{y}^S)$$

so that we have an asymptotic recovery of the reflectivity.

This calculation in media with smooth background velocity was carried out by Baylikn in the 80's. It is presented in the book of Bleistein, Cohen and Stockwell (Springer 2001).

By an elementary calculation we find that the multiplier M is proportional to

$$M(\mathbf{x}, \mathbf{y}^S, \omega) \sim \mathbf{n}(\mathbf{x}) \cdot (\mathbf{x} - \mathbf{y}^S)$$

where $\mathbf{n}(\mathbf{x})$ is the unit "outward" normal to array at $\mathbf{x}$. By outward we mean that it is pointing in the direction exterior to the region of the reflectors with reflectivity ρ.

This is a rather simple form for the multiplier, which suggests that there should be a more direct and perhaps more general and less computational way to get this asymptotic consistency of the migration imaging functional. In Lecture III we will see that this is indeed the case. There is a simpler and more general way to analyze the large array, large bandwidth behavior of backpropagation (migration) imaging functionals.

Note also that the multiplier M must be consistent with the least squares multiplier $A^H A$ (Lecture I), which we have dropped in travel time migration and in back propagation.

Resolution theory, use of the Kirchhoff-Helmholtz identities

Green's identity

Let D be a closed and bounded region with smooth boundary ∂D and let $c(\mathbf{x})$ be a given speed of propagation that is uniform outside a subregion interior to D. The time harmonic, outgoing Green's function and its conjugate satisfy

$$\Delta_{\mathrm{x}}\widehat{G}(\omega,\mathbf{x},\mathbf{y}_2) + \frac{\omega^2}{c^2(\mathbf{x})}\widehat{G}(\omega,\mathbf{y},\mathbf{y}_2) = -\delta(\mathbf{x}-\mathbf{y}_2)\,,$$
$$\Delta_{\mathrm{x}}\overline{\widehat{G}}(\omega,\mathbf{x},\mathbf{y}_1) + \frac{\omega^2}{c^2(\mathbf{x})}\overline{\widehat{G}}(\omega,\mathbf{x},\mathbf{y}_1) = -\delta(\mathbf{x}-\mathbf{y}_1)\,.$$

We multiply the first equation by $\overline{\widehat{G}}(\omega,\mathbf{x},\mathbf{y}_1)$ and subtract the second equation multiplied by $\widehat{G}(\omega,\mathbf{x},\mathbf{y}_2)$:

$$\nabla_{\mathrm{y}}\cdot[\overline{\widehat{G}}(\omega,\mathbf{x},\mathbf{y}_1)\nabla_{\mathrm{x}}\widehat{G}(\omega,\mathbf{x},\mathbf{y}_2) - \widehat{G}(\omega,\mathbf{x},\mathbf{y}_2)\nabla_{\mathrm{x}}\overline{\widehat{G}}(\omega,\mathbf{x},\mathbf{y}_1)]$$
$$= \widehat{G}(\omega,\mathbf{x},\mathbf{y}_2)\delta(\mathbf{x}-\mathbf{y}_1) - \overline{\widehat{G}}(\omega,\mathbf{x},\mathbf{y}_1)\delta(\mathbf{x}-\mathbf{y}_2)$$
$$= \widehat{G}(\omega,\mathbf{x}_1,\mathbf{y}_2)\delta(\mathbf{x}-\mathbf{y}_1) - \overline{\widehat{G}}(\omega,\mathbf{y}_1,\mathbf{y}_2)\delta(\mathbf{x}-\mathbf{y}_2)\,,$$

where we have used the reciprocity property $\widehat{G}(\omega,\mathbf{y}_2,\mathbf{y}_1) = \widehat{G}(\omega,\mathbf{y}_1,\mathbf{y}_2)$. Integrate over ∂D and use the divergence theorem:

$$\int_{\partial D}\mathbf{n}(\mathbf{x})\cdot[\overline{\widehat{G}}(\omega,\mathbf{x},\mathbf{y}_1)\nabla_{\mathrm{x}}\widehat{G}(\omega,\mathbf{x},\mathbf{y}_2) - \widehat{G}(\omega,\mathbf{x},\mathbf{y}_2)\nabla_{\mathrm{x}}\overline{\widehat{G}}(\omega,\mathbf{x},\mathbf{y}_1)]dS(\mathbf{x})$$
$$= \widehat{G}(\omega,\mathbf{y}_1,\mathbf{y}_2) - \overline{\widehat{G}}(\omega,\mathbf{y}_1,\mathbf{y}_2)\,,$$

where $\mathbf{n}(\mathbf{x})$ is the unit outward normal to ∂D.

The Sommerfeld radiation condition

For the time harmonic Green's function $\widehat{G}(\omega,\mathbf{x},\mathbf{y})$ this condition is

$$|\mathbf{x}|(\frac{\mathbf{x}}{|\mathbf{x}|}\cdot\nabla_{\mathbf{x}}-\frac{i\omega}{c_0})\widehat{G}(\omega,\mathbf{x},\mathbf{y})\to 0$$

as $|\mathbf{x}|\to\infty$. In a homogeneous medium where

$$\widehat{G}(\omega,\mathbf{x},\mathbf{y})=\widehat{G}_0(\omega,\mathbf{x},\mathbf{y})=\frac{e^{i(\omega/c_0)|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}$$

this simply means that for $|\mathbf{x}-\mathbf{y}|\to\infty$

$$\nabla_{\mathbf{x}}\widehat{G}_0(\omega,\mathbf{x},\mathbf{y})\approx\frac{i\omega}{c_0}\frac{\mathbf{x}-\mathbf{y}}{|\mathbf{x}-\mathbf{y}|}\widehat{G}_0(\omega,\mathbf{x},\mathbf{y})$$

For a smooth background velocity the high frequency (WKB) approximation of the Green's function is

$$\widehat{G}(\omega,\mathbf{x},\mathbf{y})\approx a(\mathbf{x},\mathbf{y})e^{i\omega\tau(\mathbf{x},\mathbf{y})}\,.$$

Here $a(\mathbf{x},\mathbf{y})$ and $\tau(\mathbf{x},\mathbf{y})$ are smooth except at $\mathbf{x}=\mathbf{y}$. The amplitude $a(\mathbf{x},\mathbf{y})$ satisfies a transport equation and the travel time $\tau(\mathbf{x},\mathbf{y})$ the eikonal equation. It is symmetric $\tau(\mathbf{x},\mathbf{y})=\tau(\mathbf{y},\mathbf{x})$ and from Fermat's principle

$$\tau(\mathbf{x},\mathbf{y})=\inf\left\{T \text{ s.t.} \exists\,(\mathbf{X}_t)_{t\in[0,T]}\in\mathcal{C}^1\,,\ \mathbf{X}_0=\mathbf{x}\,,\ \mathbf{X}_T=\mathbf{y}\,,\ |\frac{d\mathbf{X}_t}{dt}|=c(\mathbf{X}_t)\right\}.$$

The radiation condition can now be written as

$$\nabla_{\mathbf{x}}\widehat{G}(\omega,\mathbf{x},\mathbf{y})\approx i\omega\nabla_{\mathbf{x}}\tau(\mathbf{x},\mathbf{y})\widehat{G}(\omega,\mathbf{x},\mathbf{y})$$

Now lets assume that we can use the radiation condition in Green's identity. We get

$$i\omega \int_{\partial D} \mathbf{n}(\mathbf{x}) \cdot (\nabla_{\mathbf{x}} \tau(\mathbf{x}, \mathbf{y}^S) + \nabla_{\mathbf{x}} \tau(\mathbf{x}, \mathbf{y})) \overline{\widehat{G}(\omega, \mathbf{x}, \mathbf{y})} \widehat{G}(\omega, \mathbf{x}, \mathbf{y}^S) dS(\mathbf{x})$$
$$= \widehat{G}(\omega, \mathbf{y}, \mathbf{y}^S) - \overline{\widehat{G}(\omega, \mathbf{y}, \mathbf{y}^S)},$$

Assume that $\mathbf{y}$ and $\mathbf{y}^S$ are near each other as they are the source location and the search point of the imaging function. Then we have

$$2i\omega \int_{\partial D} \mathbf{n}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} \tau(\mathbf{x}, \mathbf{y}^S) \overline{\widehat{G}(\omega, \mathbf{x}, \mathbf{y})} \widehat{G}(\omega, \mathbf{x}, \mathbf{y}^S) dS(\mathbf{x})$$
$$= \widehat{G}(\omega, \mathbf{y}, \mathbf{y}^S) - \overline{\widehat{G}(\omega, \mathbf{y}, \mathbf{y}^S)},$$

Connection with imaging and time reversal

Clearly

$$I^{WBP}(\mathbf{y}^S) = \int_{\partial D} \mathbf{n}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} \tau(\mathbf{x}, \mathbf{y}^S) \overline{\widehat{G}(\omega, \mathbf{x}, \mathbf{y})} \widehat{G}(\omega, \mathbf{x}, \mathbf{y}^S) dS(\mathbf{x})$$

is a weighted imaging functional with back propagation to a search point $\mathbf{y}^S$, when there is a point source at $\mathbf{y}$ and the array is the "full aperture" boundary ∂D.

From the KH identity, assuming that it can used, we have that

$$I^{WBP}(\mathbf{y}^S) \approx \frac{1}{2i\omega}(\widehat{G}(\omega, \mathbf{y}, \mathbf{y}^S) - \overline{\widehat{G}(\omega, \mathbf{y}, \mathbf{y}^S)}).$$

This approximate identity is quite general regarding the background medium, which can be rough and even random. But (i)
the array must be in a homogeneous medium and far from the scattering background. The source to be "imaged" can be in the scattering region. We put quotations on imaged because if the medium is rough and random it will not be known and so $I^{WBP}(\mathbf{y}^S)$ is a time reversal field function, not an imaging functional. If the background is variable but known then it is an imaging functional. In addition (ii)
the array must sufficiently far so that the radiation condition applies.

In a smooth background at high frequencies we have

$$\widehat{G}(\omega, \mathbf{x}, \mathbf{y}) \approx a(\mathbf{x}, \mathbf{y}) e^{i\omega\tau(\mathbf{x},\mathbf{y})}$$

and lets assume that the amplitude a is real. Then

$$\begin{aligned} I^{WBP}(\mathbf{y}^S) &\approx \frac{1}{2i\omega}(\widehat{G}(\omega, \mathbf{y}, \mathbf{y}^S) - \overline{\widehat{G}(\omega, \mathbf{y}, \mathbf{y}^S)}) \\ &\approx \frac{1}{\omega} a(\mathbf{y}, \mathbf{y}^S) \sin(\omega\tau(\mathbf{y}, \mathbf{y}^S)) \end{aligned}$$

From the first zero of the sine function we get a resolution limit, the Rayleigh resolution limit: $\omega\tau(\mathbf{y}, \mathbf{y}^*) = \pi$. If c_0 is a reference background speed then

$$c_0\tau(\mathbf{y}, \mathbf{y}^*) = \frac{\pi c_0}{\omega} = \frac{\lambda_0}{2},$$

which is an appropriate generalization of the "half wavelength" far field (high frequency) resolution limit for full aperture imaging.

In the high frequency, far field regime we can symmetrize a large planar array about the source and make the array imaging functional (the psf) look approximately like a full aperture functional.

This argument leads to the approximation

$$\int_{|\omega-\omega_0|\le B/2} d\omega \widehat{f}_B(\omega-\omega_0) \int_A \mathbf{n}(\mathbf{x})\cdot\nabla_{\mathbf{x}}\tau(\mathbf{x},\mathbf{y}^S)\overline{\widehat{G}(\omega,\mathbf{x},\mathbf{y})}\widehat{G}(\omega,\mathbf{x},\mathbf{y}^S)dS(\mathbf{x})$$

$$\approx a(\mathbf{y},\mathbf{y}^S)\int_{|\omega-\omega_0|\le B/2} d\omega \frac{\widehat{f}_B(\omega-\omega_0)}{4i\omega} e^{i\omega\tau(\mathbf{y},\mathbf{y}^S)}$$

We have used here the fact that the pulse is a real function and the omega integration extends to negative frequencies with $\overline{\widehat{f}_B(\omega)} = \widehat{f}_B(-\omega)$.

For a suitably chosen probing pulse we see from this approximation that

$$\int_{|\omega-\omega_0|\leq B/2} d\omega \widehat{f}_B(\omega-\omega_0) \int_A \mathbf{n}(\mathbf{x})\cdot\nabla_{\mathbf{x}}\tau(\mathbf{x},\mathbf{y}^S)\overline{\widehat{G}(\omega,\mathbf{x},\mathbf{y})}\widehat{G}(\omega,\mathbf{x},\mathbf{y}^S)dS(\mathbf{x})$$

$$\approx a(\mathbf{y},\mathbf{y}^S)\delta(\tau(\mathbf{y},\mathbf{y}^S)),$$

as $B \to \infty$ and $A \to R^2$.

This is a generalization of the resolution result mentioned earlier in connection with a change of variables in a homogeneous medium.

Summary of results in resolution theory

- Time-harmonic, paraxial, $\lambda \ll a \ll L$:
  Cross range $(\lambda L)/a$, Range $\lambda(L/a)^2$.

- Broadband: Range c_0/B (arrival time resolution). Cross range resolution is still $(\lambda L)/a$.

- Full aperture: $\lambda/2$

- Large aperture, large bandwidth arrays: exact recovery with suitably weighted migration or back propagation functional.

Noise sources and correlations. Open media and cavities. Velocity estimation and imaging with distributed sensors

Based on a paper with Josselin Garnier that can be obtained from `http://math.stanford.edu/~papanico`

The simplest way randomness can enter is from sources that are (i) randomly distributed and (ii) are stationary random processes in time.

In this case the signals recorded at sensors located at $\{\mathbf{x}_j\}$, usually distributed over some region, are themselves stationary in time random processes. What information can possibly be in these signals? Can we image with them?

First some definitions:
1. Random process $\nu(t), t \in R$ (or $\mu(\mathbf{x}), \in R^d, d > 1$) are stationary if $\nu(t_1), \nu(t_2), \ldots, \nu(t_M)$ has the same joint law as $\nu(t_1+h), \nu(t_2+h), \ldots, \nu(t_M+h)$ for any set of points $\{t_j\}$ and any h. In this case $E\{\nu(t)\}$ is a constant which we take as zero.
2. The correlation $C(\tau) = E\{\nu(t)\nu(t+\tau)\}$ is a function of the lag τ only. Assuming that C is an integrable function, its Fourier transform $\widehat{C}(\omega) = \int e^{i\omega t} C(t) dt$ is always non-negative or more generally a measure (Bohner's theorem).

Let $u(t, \mathbf{x}_1)$ and $u(t, \mathbf{x}_2)$ denote the time-dependent wave fields recorded by two sensors at $\mathbf{x}_1$ and $\mathbf{x}_2$. Their empirical cross correlation function over the time interval $[0, T]$ with time lag τ is given by

$$C_T(\tau, \mathbf{x}_1, \mathbf{x}_2) = \frac{1}{T} \int_0^T u(t, \mathbf{x}_1) u(t + \tau, \mathbf{x}_2) dt \,.$$

In a homogeneous medium, if the source of the waves is a space-time stationary random field that is also delta correlated in space and time then we will show that

$$\frac{\partial}{\partial \tau} C_T(\tau, \mathbf{x}_1, \mathbf{x}_2) \simeq G(\tau, \mathbf{x}_1, \mathbf{x}_2) - G(-\tau, \mathbf{x}_1, \mathbf{x}_2) \,,$$

where G is the Green's function.

This approximate equality holds for T sufficiently large and provided some limiting absorption is introduced to regularize the integral. The main point here is that the time-symmetrized Green's function can be obtained from the cross correlation if there is enough source diversity. In this case the wave field at any sensor is equipartitioned, in the sense that it is a superposition of uncorrelated plane waves of all directions. We can recover in particular the travel time $\tau(\mathbf{x}_1, \mathbf{x}_2)$ from the singular support of the cross correlation.

Wave equation with noise sources

We consider solutions u of the wave equation in a three dimensional inhomogeneous (possibly random: $c(\mathbf{x})$ random) medium:

$$\frac{1}{c^2(\mathbf{x})}\frac{\partial^2 u}{\partial t^2} - \Delta_{\mathbf{x}} u = n^\varepsilon(t, \mathbf{x}) \, .$$

The term $n^\varepsilon(t, \mathbf{x})$ models a random distribution of noise sources. It is a zero-mean stationary (in time) Gaussian process with autocorrelation function

$$\langle n^\varepsilon(t_1, \mathbf{y}_1) n^\varepsilon(t_2, \mathbf{y}_2) \rangle = F^\varepsilon(t_2 - t_1)\Gamma(\mathbf{y}_1, \mathbf{y}_2) \, .$$

Here $\langle \cdot \rangle$ stands for statistical average with respect to the distribution of the noise sources.

We assume that the decoherence time of the noise sources is much smaller than typical travel times between sensors. If we denote with ε the (small) ratio of these two time scales, we can then write the time correlation function F^ε in the form

$$F^\varepsilon(t_2 - t_1) = F(\frac{t_2 - t_1}{\varepsilon}) \, ,$$

where t_1 and t_2 are scaled relative to typical sensor travel times.

The Fourier transform $\widehat{F}^{\varepsilon}$ of the time correlation function is a nonnegative, even, real-valued function. It is proportional to the power spectral density of the sources:

$$\widehat{F}^{\varepsilon}(\omega) = \varepsilon \widehat{F}(\varepsilon\omega)\,,$$

where the Fourier transform is defined by

$$\widehat{F}(\omega) = \int F(t) e^{i\omega t} dt\,.$$

The spatial distribution of the noise sources is characterized by the autocovariance function Γ. It is the kernel of a symmetric nonnegative definite operator. For simplicity, we will assume that the process n is delta-correlated in space:

$$\Gamma(\mathbf{y}_1, \mathbf{y}_2) = \theta(\mathbf{y}_1)\delta(\mathbf{y}_1 - \mathbf{y}_2)\,,$$

where θ characterizes the spatial support of the sources.

The stationary solution of the wave equation has the integral representation

$$\begin{aligned} u(t,\mathbf{x}) &= \int\int_{-\infty}^{t} n^{\varepsilon}(s,\mathbf{y})G(t-s,\mathbf{x},\mathbf{y})ds d\mathbf{y} \\ &= \int\int n^{\varepsilon}(t-s,\mathbf{y})G(s,\mathbf{x},\mathbf{y})ds d\mathbf{y}\,, \end{aligned}$$

where $G(t,\mathbf{x},\mathbf{y})$ is the time-dependent Green's function. It is the fundamental solution of the wave equation

$$\frac{1}{c^2(\mathbf{x})}\frac{\partial^2 G}{\partial t^2} - \Delta_{\mathbf{x}} G = \delta(t)\delta(\mathbf{x}-\mathbf{y})\,,$$

starting from $G(0,\mathbf{x},\mathbf{y}) = \partial_t G(0,\mathbf{x},\mathbf{y}) = 0$ (and continued on the negative time axis by $G(t,\mathbf{x},\mathbf{y}) = 0\ \forall t \leq 0$).

The empirical cross correlation and its mean

The empirical cross correlation of the signals recorded at $\mathbf{x}_1$ and $\mathbf{x}_2$ for an integration time T is

$$C_T(\tau, \mathbf{x}_1, \mathbf{x}_2) = \frac{1}{T} \int_0^T u(t, \mathbf{x}_1) u(t+\tau, \mathbf{x}_2) dt \,.$$

It is a statistically stable quantity, in the sense that for a large integration time T, C_T is independent of the realization of the noise sources.

The expectation of C_T (with respect to the distribution of the sources) is independent of T:

$$\langle C_T(\tau, \mathbf{x}_1, \mathbf{x}_2) \rangle = C^{(1)}(\tau, \mathbf{x}_1, \mathbf{x}_2) \,,$$

where $C^{(1)}$ is given by

$$C^{(1)}(\tau, \mathbf{x}_1, \mathbf{x}_2) = \int d\mathbf{y} \int ds ds' G(s, \mathbf{x}_1, \mathbf{y}) G(\tau+s+s', \mathbf{x}_2, \mathbf{y}) F^{\varepsilon}(s') \theta(\mathbf{y}) \,,$$

Fourier form and self-averaging of the cross correlation

In the frequency domain the cross correlation is given by

$$C^{(1)}(\tau,\mathbf{x}_1,\mathbf{x}_2) = \int d\mathbf{y} \int d\omega \overline{\widehat{G}}(\omega,\mathbf{x}_1,\mathbf{y})\widehat{G}(\omega,\mathbf{x}_2,\mathbf{y})\widehat{F}^{\varepsilon}(\omega)e^{-i\omega\tau}\theta(\mathbf{y})\,.$$

The empirical cross correlation C_T is a self-averaging quantity:

$$C_T(\tau,\mathbf{x}_1,\mathbf{x}_2) \stackrel{T\to\infty}{\longrightarrow} C^{(1)}(\tau,\mathbf{x}_1,\mathbf{x}_2)\,,$$

in probability with respect to the distribution of the sources.

More precisely, the fluctuations of C_T around its mean value $C^{(1)}$ are of order $T^{-1/2}$ for T large compared to the decoherence time of the sources.

We derive the relation between the cross correlation and the Green's function when the medium is homogeneous with background velocity c_0 and the source distribution extends over all space, i.e. $\theta \equiv 1$.

In this case the signal amplitude diverges because the contributions from noise sources far away from the sensors are not damped. For a well-posed formulation we need to introduce some dissipation, so we consider the solution u of the damped wave equation:

$$\frac{1}{c_0^2}\left(\frac{1}{T_a} + \frac{\partial}{\partial t}\right)^2 u - \Delta_{\mathbf{x}} u = n^{\varepsilon}(t, \mathbf{x}) \, .$$

Cross correlation identity in a homogeneous medium

In a three-dimensional medium with dissipation and with noise source distribution extending over all space $\theta \equiv 1$

$$\frac{\partial}{\partial \tau} C^{(1)}(\tau, \mathbf{x}_1, \mathbf{x}_2) = -\frac{c_0^2 T_a}{4} e^{-\frac{|\mathbf{x}_1 - \mathbf{x}_2|}{c_0 T_a}} \left[F^\varepsilon * G(\tau, \mathbf{x}_1, \mathbf{x}_2) - F^\varepsilon * G(-\tau, \mathbf{x}_1, \mathbf{x}_2) \right],$$

where $*$ denotes convolution in τ and G is the Green's function of the homogeneous medium without dissipation:

$$G(t, \mathbf{x}_1, \mathbf{x}_2) = \frac{1}{4\pi |\mathbf{x}_1 - \mathbf{x}_2|} \delta\left(t - \frac{|\mathbf{x}_1 - \mathbf{x}_2|}{c_0}\right).$$

Estimating travel times from cross correlations

If the decoherence time of the sources is much shorter than the travel time (i.e., $\varepsilon \ll 1$), then F^ε behaves like a Dirac distribution and we have

$$\frac{\partial}{\partial \tau} C^{(1)}(\tau, \mathbf{x}_1, \mathbf{x}_2) \simeq e^{-\frac{|\mathbf{x}_1 - \mathbf{x}_2|}{c_0 T a}} \left[G(\tau, \mathbf{x}_1, \mathbf{x}_2) - G(-\tau, \mathbf{x}_1, \mathbf{x}_2) \right],$$

up to a multiplicative constant.

It is therefore possible to estimate the travel time $\tau(\mathbf{x}_1, \mathbf{x}_2) = |\mathbf{x}_1 - \mathbf{x}_2|/c_0$ between $\mathbf{x}_1$ and $\mathbf{x}_2$ from the cross correlation, with an accuracy of the order of the decoherence time of the noise sources.

On the extraction of the travel time from cross correlations

The cross correlation is closely related to the symmetrized Green's function from $\mathbf{x}_1$ to $\mathbf{x}_2$ not only for a homogeneous medium but also for inhomogeneous media.

One can give a simple and rigorous proof for an open inhomogeneous medium in the case in which the noise sources are located on the surface of a sphere that encloses both the inhomogeneous region and the sensors, located at $\mathbf{x}_1$ and $\mathbf{x}_2$.

The proof is based on an approximate identity that follows from Green's identity and the Sommerfeld radiation condition. This approximate identity is none other than the Helmholtz-Kirchhoff integral theorem of Lecture III.

If the sensors are at known locations $\{\mathbf{x}_j\}, j = 1, 2, \ldots, N$ and are suitably distributed over a region whose speed of propagation $c(\mathbf{x})$ is unknown, then this speed can be estimated from the travel time $\{\tau(\mathbf{x}_i, \mathbf{x}_j)\}$. There are tomographic algorithms for doing this estimation. The accuracy and robustness of the resulting estimate $\widehat{c}(\mathbf{x})$ will depend on (i) the accuracy of the travel time estimates, (ii) the topology of the sensor network, and (iii) the properties of the ambient noise sources, which are also unknown.

The estimation of the surface wave velocity in Southern California from seismic noise correlations over some 150 seismic stations was a breakthrough in 2005 when it was successfully done by Sabra, Gerstoft, Roux, and Kuperman (Surface wave tomography from microseisms in Southern California, *Geophys. Res. Lett.* **32** L14311)

Derivation of the cross correlation identity

The Green's function of the homogeneous medium with dissipation is:

$$G_a(t, \mathbf{x}_1, \mathbf{x}_2) = G(t, \mathbf{x}_1, \mathbf{x}_2) e^{-\frac{t}{T_a}}.$$

The cross correlation function is given by

$$C^{(1)}(\tau, \mathbf{x}_1, \mathbf{x}_2) = \int d\mathbf{y} \int ds ds' G_a(s, \mathbf{x}_1, \mathbf{y}) G_a(\tau + s + s', \mathbf{x}_2, \mathbf{y}) F^\varepsilon(s').$$

Integrating in s and s' gives

$$C^{(1)}(\tau, \mathbf{x}_1, \mathbf{x}_2) = \int \frac{d\mathbf{y}}{16\pi^2 |\mathbf{x}_1 - \mathbf{y}| \, |\mathbf{x}_2 - \mathbf{y}|} e^{-\frac{|\mathbf{x}_1 - \mathbf{y}| + |\mathbf{x}_2 - \mathbf{y}|}{c_0 T_a}} F^\varepsilon(\tau - \frac{|\mathbf{x}_1 - \mathbf{y}| - |\mathbf{x}_2 - \mathbf{y}|}{c_0}).$$

We parameterize the locations of the sensors by $\mathbf{x}_1 = (h, 0, 0)$ and $\mathbf{x}_2 = (-h, 0, 0)$, where $h > 0$, and we use the change of variables for $\mathbf{y} = (x, y, z)$:

$$\begin{cases} x = h \sin\theta \cosh\phi \,, & \phi \in (0, \infty)\,, \\ y = h \cos\theta \sinh\phi \cos\psi \,, & \theta \in (-\pi/2, \pi/2)\,, \\ z = h \cos\theta \sinh\phi \sin\psi \,, & \psi \in (0, 2\pi)\,, \end{cases}$$

whose Jacobian is $J = h^3 \cos\theta \sinh\phi (\cosh^2\psi - \sin^2\theta)$. Using the fact that $|\mathbf{x}_1 - \mathbf{y}| = h(\cosh\phi - \sin\theta)$ and $|\mathbf{x}_2 - \mathbf{y}| = h(\cosh\phi + \sin\theta)$, we get

$$C^{(1)}(\tau, \mathbf{x}_1, \mathbf{x}_2) = \frac{h}{8\pi} \int_0^\infty d\phi \sinh\phi \int_{-\pi/2}^{\pi/2} d\theta \cos\theta e^{-\frac{2h\cosh\phi}{c_0 T_a}} F^\varepsilon(\tau + \frac{2h \sin\theta}{c_0}).$$

Derivation continued

After the new change of variables $u = h\cosh\phi$ and $s = (2h/c_0)\sin\theta$, we obtain

$$C^{(1)}(\tau, \mathbf{x}_1, \mathbf{x}_2) = \frac{c_0^2 T_a}{32\pi h} e^{-\frac{2h}{c_0 T_a}} \int_{-2h/c_0}^{2h/c_0} F^\varepsilon(\tau + s) ds \, .$$

By differentiating in τ, we get

$$\frac{\partial}{\partial \tau} C^{(1)}(\tau, \mathbf{x}_1, \mathbf{x}_2) = \frac{c_0^2 T_a}{32\pi h} e^{-\frac{2h}{c_0 T_a}} [F^\varepsilon(\tau + \frac{2h}{c_0}) - F^\varepsilon(\tau - \frac{2h}{c_0})] \, ,$$

which is the desired result since $|\mathbf{x}_1 - \mathbf{x}_2| = 2h$.

Let us assume that the medium is homogeneous with background velocity c_e outside the ball $B(\mathbf{0}, r)$ with center $\mathbf{0}$ and radius r. Then, for any $\mathbf{x}_1, \mathbf{x}_2 \in B(\mathbf{0}, r)$ we have for $L \gg r$ the KH identity:

$$\widehat{G}(\omega, \mathbf{x}_1, \mathbf{x}_2) - \overline{\widehat{G}}(\omega, \mathbf{x}_1, \mathbf{x}_2) = \frac{2i\omega}{c_e} \int_{\partial B(\mathbf{0}, L)} \overline{\widehat{G}}(\omega, \mathbf{x}_1, \mathbf{y}) \widehat{G}(\omega, \mathbf{x}_2, \mathbf{y}) dS(\mathbf{y}) \, .$$

We also assume that the sources are localized with a uniform density on the sphere $\partial B(\mathbf{0}, L)$ with center $\mathbf{0}$ and radius L.

If $L \gg r$, then for any $\mathbf{x}_1, \mathbf{x}_2 \in B(\mathbf{0}, r)$

$$\frac{\partial}{\partial \tau} C^{(1)}(\tau, \mathbf{x}_1, \mathbf{x}_2) = -F^{\varepsilon} * G(\tau, \mathbf{x}_1, \mathbf{x}_2) + F^{\varepsilon} * G(-\tau, \mathbf{x}_1, \mathbf{x}_2) \, ,$$

up to a multiplicative factor. Here $*$ stands for convolution in τ.

Directional sources

What if the ambient noise sources are not distributed evenly about the sensors? What if the wave fields recorded at the sensors have a dominant orientation instead of being equdistributed in all directions?

In such cases we cannot expect to be able to recover the full (symmetrized) Green's function between the sensors. At best we can recover the travel time $\tau(\mathbf{x}_1, \mathbf{x}_2)$ if the line (the ray) connecting the two sensors continues into the source region. This is done using the stationary phase method and is discussed further in Garnier's lecture (Thursday afternoon).

What about imaging reflectors with passive sensor networks using ambient noise sources? This can be done using suitable fourth order cross correlations.

Concluding remarks on noise cross correlations

- Cross correlations can be used effectively in closed environments with limited ambient noise source diversity but enhancing by multiple wall reflections (ergodic cavities)
- It is also possible that a scattering medium can enhance ambient noise source diversity and make the estimation of background velocities feasible
- But there is a limitation in how strongly scattering the medium can be: the transport mean free path must be long compared to the distance between sensors (to preserve coherence) but short compared to the distance between noise sources and sensors
- In a scattering medium, the transport mean free path is a rough measure of how far waves have to propagate before they lose their coherence and wave energy diffuses isotropically

Imaging in Random Media

George Papanicolaou
Department of Mathematics, Stanford University
http://georgep.stanford.edu

$Id: NSF-CBMS, CAAM Rice, May 12-16, 2008 $

The singular value decomposition, in detection and imaging

Let $\widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega)$ be the array impulse response matrix over the bandwidth $\omega_0 - B/2 < \omega < \omega_0 + B/2$, with N_s sources and N_r receivers which we will assume are collocated and $N_s = N_r = N$.

If $\widehat{f}(\omega) = (\widehat{f}(\mathbf{x}_s, \omega))$ is a vector of illuminations in the frequency domain, then

$$\widehat{P}_f(\omega) = \left(\sum_{\mathbf{x}_s} \widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega) \widehat{f}(\mathbf{x}_s, \omega) \right)$$

is the vector of received signals at the array, in the frequency domain. The total power of these signals is $\mathcal{P}_{tot}(f) = \int d\omega ||\widehat{\Pi}(\omega) \widehat{f}(\omega)||^2$

Problem: Find $\mathcal{P} = \max_f \mathcal{P}_{tot}(f)$ with $||f||^2 = \int d\omega ||\widehat{f}(\omega)||^2 = 1$

This problem of optimal illumination for received power, that is, for detection, is solved using the SVD of $\widehat{\Pi}$. We assume that we have a fixed bandwidth of size B.

The SVD of the array response matrix

The array impulse response matrix is symmetric but not hermitian $\widehat{\Pi}^T(\omega) = \widehat{\Pi}(\omega)$. Let $\hat{\mathbf{v}}_j$ and $\hat{\mathbf{u}}_j$ be its right and left singular vectors respectively. Then its singular value decomposition is

$$\widehat{\Pi}(\omega) = \sum_{j=1}^{p} \sigma_j(\omega)\hat{\mathbf{u}}_j(\omega)\hat{\mathbf{v}}_j^*(\omega)$$

Here $p \leq N$ is the rank of $\widehat{\Pi}$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p > 0$.

Suppose that $\omega^* = \text{argmax}\sigma_1(\omega)$ over the bandwidth and let

$$\widehat{f}(\omega) = \frac{1}{2\delta}\hat{\mathbf{v}}_1(\omega^*) \ , \quad \omega \in [\omega^* - \delta, \omega^* + \delta]$$

and zero outside this interval. Then for this illumination f we have that $\mathcal{P}_{tot}(f) \to \mathcal{P} = \sigma_1^2(\omega^*)$ as $\delta \to 0$. The optimal illumination is a narrow band signal proportional to $\hat{\mathbf{v}}_1(\omega^*)$.

Consider the following iterative process (experiment). It is done in the time domain but we describe it frequency by frequency:

1. Start with illumination $\widehat{f}$. The received signal at the array is $\widehat{\Pi}\widehat{f}$

2. Use the time reversed field as illumination. The received field is $\widehat{\Pi}\overline{\widehat{\Pi}\widehat{f}}$

3. Repeat these two steps n times.

The field received at the array has the form

$$\widehat{K}^n(\omega)\overline{\widehat{f}(\omega)}\ , \quad \widehat{K}(\omega) = \widehat{\Pi}(\omega)\overline{\widehat{\Pi(\omega)}}$$

where $\widehat{K}(\omega)$ is the time reversal operator. It is hermitian and positive definite for each frequency, and its eigenvalues are the squares of the singular values. Therefore for large n we have

$$\widehat{K}^n(\omega)\overline{\widehat{f}(\omega)} \approx \sigma_1^{2n}(\omega)\hat{\mathbf{v}}_1(\omega)\hat{\mathbf{v}}_1^*(\omega)\overline{\widehat{f}(\omega)}$$

Iterative time reversal, time domain

In the time domain the signal received at the array after a large number n of iterative TR has the approximate form

$$\int d\omega e^{-i\omega t}\sigma_1^{2n}(\omega)\hat{\mathbf{v}}_1(\omega)\hat{\mathbf{v}}_1^*(\omega)\overline{\hat{f}(\omega)}$$

By the Laplace asymptotic method it can be further approximated, up to a constant, by

$$e^{-i\omega^* t}\sigma_1^{2n}(\omega^*)\hat{\mathbf{v}}_1(\omega^*)\hat{\mathbf{v}}_1^*(\omega^*)\overline{\hat{f}(\omega^*)}$$

which is a time harmonic signal at the frequency where $\sigma_1(\omega)$ takes it maximum value.

With ITR we can get $\hat{\mathbf{v}}_1(\omega)$ directly from the physical experiment without doing the SVD. This however requires some special adjustments in order to get it over the full bandwidth. The other singular vectors can also be obtained with ITR.

But why are we interested in the SVD of the response matrix and ITR, which is a physical way of getting the SVD?

Point scatterer models

Let M point scatterers be at $\mathbf{y}_j, j = 1, 2, ..., M$. Point scatterers means that the reflectivity is

$$\rho(\mathbf{z}) = \sum_{j=1}^{M} \rho_j \mathbf{1}_{|\mathbf{z}-\mathbf{y}_j| \le \delta_j},$$

with the radii δ_j small compared to wavelengths. In this case the impulse response matrix in the Born approximation is

$$\widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega) = \sum_{j=1}^{M} \xi_j(\omega) \widehat{G}(\mathbf{x}_s, \mathbf{y}_j, \omega) \widehat{G}(\mathbf{x}_r, \mathbf{y}_j, \omega)$$

The scattering amplitudes $\xi_j(\omega)$ depend on the reflectivities and radii (or shape, in general), and on the frequency. Define the array vector Green's function

$$\widehat{g}(\mathbf{y}, \omega) = (\widehat{G}(\mathbf{x}_r, \mathbf{y}, \omega))$$

Then the array impulse response matrix has the form

$$\widehat{\Pi}(\omega) = \sum_{j=1}^{M} \xi_j(\omega) \widehat{g}(\mathbf{y}_j, \omega) \widehat{g}^T(\mathbf{y}_j, \omega)$$

The array vectors $\{\widehat{g}(\mathbf{y}_j, \omega)\}$ are not of course orthogonal in general. But

$$\widehat{g}^*(\mathbf{y}_j, \omega)\widehat{g}(\mathbf{y}_l, \omega) = \sum_{\mathbf{x}_r} \overline{\widehat{G}(\mathbf{x}_r, \mathbf{y}_j, \omega)}\widehat{G}(\mathbf{x}_r, \mathbf{y}_l, \omega)$$

is exactly the basic quantity that arises in imaging and time reversal, and whose behavior we have analyzed in Lectures II-III. We know that if the distance $|\mathbf{y}_j - \mathbf{y}_l|$ is large compared to the resolution limit of the array at this frequency, then these array vectors are approximately orthogonal

$$\widehat{g}^*(\mathbf{y}_j, \omega)\widehat{g}(\mathbf{y}_l, \omega) \approx ||\widehat{g}(\mathbf{y}_j, \omega)||^2 \delta_{jl}$$

In any case we may assume that the $\{\widehat{g}(\mathbf{y}_j, \omega)\}$ are linearly independent.

In the well separated case the array impulse response matrix is in SVD form.

Well separated scatterers, continued

In the well separated case we have

$$\widehat{\Pi}\overline{\widehat{g}(\mathbf{y}_l)} = \sum_{j=1}^{M} \xi_j \widehat{g}(\mathbf{y}_j)\widehat{g}^T(\mathbf{y}_j)\overline{\widehat{g}(\mathbf{y}_l)} \approx \xi_l ||\widehat{g}(\mathbf{y}_l)||^2 \widehat{g}(\mathbf{y}_l)$$

Therefore assuming that the ξ_l are positive and that $\xi_l||\widehat{g}(\mathbf{y}_l)||^2$ are arranged in decreasing order we have

$$\widehat{\mathbf{v}}_l = \frac{\overline{\widehat{g}(\mathbf{y}_l)}}{||\widehat{g}(\mathbf{y}_l)||}, \quad \widehat{\mathbf{u}}_l = \overline{\widehat{\mathbf{v}}_l}, \quad \sigma_l = \xi_l ||\widehat{g}(\mathbf{y}_l)||^2$$

We conclude that the rank of the SVD can be associated uniquely with the number of small scatterers, even of they are not well separated, up to some special configurations.

We now look at time reversal and imaging with the SVD

With illumination f the time reversal field at $\mathbf{y}^S$ is

$$\Gamma_f^{TR}(\mathbf{y}^S) = \int d\omega \widehat{g}^T(\mathbf{y}^S, \omega)\overline{\widehat{\Pi}(\omega)\widehat{f}(\omega)}$$

When $\widehat{f} = \widehat{\mathbf{v}}_l$ then

$$\Gamma_l^{TR}(\mathbf{y}^S) = \int d\omega \sigma_l(\omega)\widehat{g}^T(\mathbf{y}^S, \omega)\overline{\widehat{\mathbf{u}}_l(\omega)}$$

and in the well separated case

$$\Gamma_l^{TR}(\mathbf{y}^S) = \int d\omega \xi_l(\omega)||\widehat{g}(\mathbf{y}_l, \omega)||\widehat{g}^T(\mathbf{y}^S, \omega)\overline{\widehat{g}(\mathbf{y}_l, \omega)}$$

What is interesting here is that by using the SVD we can selectively do time reversal to the l-th scatterer. And by using iterative time reversal we can do this completely in hardware, without doing a numerical SVD. There are advantages to this when SNR issues are important.

In order to image we have to back propagate in a known medium, which we take as a homogeneous one and let

$$\widehat{g}_0(\mathbf{y},\omega) = (\widehat{G}_0(\mathbf{x}_r,\mathbf{y},\omega))$$

The Kirchhoff migration functional in then given by

$$I^{KM}(\mathbf{y}^S) = \int d\omega \widehat{g}_0^T(\mathbf{y}^S,\omega)\overline{\widehat{\Pi}(\omega)}\widehat{g}_0(\mathbf{y}^S,\omega)$$

$$= \int d\omega \sum_{j=1}^{p} \sigma_j(\omega)\widehat{g}_0^T(\mathbf{y}^S,\omega)\overline{\widehat{\mathbf{u}}_j(\omega)}\widehat{\mathbf{v}}_j^T(\omega)\widehat{g}_0(\mathbf{y}^S,\omega)$$

In the case of well separated scatterers we have

$$I^{KM}(\mathbf{y}^S) = \int d\omega \sum_{j=1}^{p} \xi_j(\omega)|\widehat{g}_0^*(\mathbf{y}^S,\omega)\widehat{g}_0(\mathbf{y}_j,\omega)|^2$$

We see now how the basic resolution theory of the source point spread function can be carried over to KM.

We see that KM imaging, which is the the unfiltered least squares imaging functional ($A^H A \approx I$), is rather strange because it illuminates a spot and then back propagates to it with the same array vector $\widehat{g}_0(\mathbf{y}^S, \omega)$. If we choose a general illumination vector f, a linear combination of the right singular vectors for example, we have

$$I^{BP}(\mathbf{y}^S; f) = \int d\omega \widehat{g}_0^T(\mathbf{y}^S, \omega)\overline{\widehat{\Pi}(\omega)}\widehat{f}(\omega)$$

If we let $\widehat{f}(\omega) = \sum d_l(\omega)\widehat{\mathbf{v}}_l(\omega)$ then

$$I^{BP}(\mathbf{y}^S; d) = \int d\omega \sum_{j=1}^{p} \sigma_j(\omega)\widehat{g}_0^T(\mathbf{y}^S, \omega)\overline{\widehat{\mathbf{u}}_j(\omega)}\widehat{\mathbf{v}}_j^T(\omega) \sum_{l=1}^{p} d_l(\omega)\overline{\widehat{\mathbf{v}}_l(\omega)}$$

$$I^{BP}(\mathbf{y}^S; d) = \int d\omega \sum_{j=1}^{p} \sigma_j(\omega) d_j(\omega)\widehat{g}_0^T(\mathbf{y}^S, \omega)\overline{\widehat{\mathbf{u}}_j(\omega)}$$

We can now look for a way to choose the weights $\{d_j(\omega)\}$ so as to optimize the image.

Let

$$\mathcal{G}_j(\omega) = \int d\mathbf{y}^S |\widehat{g}_0^T(\mathbf{y}^S,\omega)\overline{\widehat{\mathbf{u}}_j(\omega)}|^2$$

with the integration over some window, and let

$$\mathcal{I}(d) = \int d\omega \sum_{j=1}^{p} \sigma_j(\omega) d_j(\omega) \mathcal{G}_j(\omega)$$

Then we can try to find weights $\{d_j(\omega)\}$ that minimize this objective function.

There is no reason to adhere to the array least squares criterion, which leads to filtered back propagation (or back projection), as a basis for imaging.

Criteria based on the quality of the image directly have many advantages, especially in random media.

Another way to introduce an optimization process using the SVD is by subspace selection. Let

$$D\left[\widehat{\Pi}(\omega);\omega\right] = \sum_{j=1}^{p} \sigma_j(\omega) d_j(\omega) \hat{\mathbf{u}}_j(\omega) \hat{\mathbf{v}}_j^*(\omega)$$

a subspace selector with weights $\{d_j(\omega)\}$. Now consider KM imaging with it instead of $\widehat{\Pi}$. We have

$$I^{KM}(\mathbf{y}^S; d) = \int d\omega \widehat{g}_0^T(\mathbf{y}^S,\omega)\overline{D\left[\widehat{\Pi}(\omega);\omega\right]}\widehat{g}_0(\mathbf{y}^S,\omega)$$

or

$$I^{KM}(\mathbf{y}^S; d) = \int d\omega \sum_{j=1}^{p} \sigma_j(\omega) d_j(\omega) (\hat{\mathbf{u}}_j^*(\omega)\widehat{g}_0(\mathbf{y}^S,\omega))^2$$

If we now integrate $|I^{KM}(\mathbf{y}^S; d)|$ over an image window we see that we get back an objective similar to the optimal illumination criterion, which we must minimize over $\{d_j(\omega)\}$.

Summary of SVD, TR and imaging methods

- The SVD of the array response matrix allows for selective focusing with TR on small scatterers

- The SVD singular vectors for TR can be computed directly with ITR without having to know the full response matrix in advance

- The SVD can be used for optimal illumination, or optimal subspace selection, for migration imaging that is based on the quality of the image itself

Imaging in random media

George Papanicolaou
Department of Mathematics, Stanford University
http://math.stanford.edu/~papanico

Lectures VII-VIII

Waves in random media: Layered media, the paraxial approximation, radiative transport

Time reversal in random media, super-resolution, statistical stability

We consider the wave equation in a random medium

$$\frac{1}{c^2(\vec{\mathbf{x}})}\frac{\partial^2 u}{\partial t^2} - \triangle u = 0 \ , \quad t > 0 \ , \quad \vec{\mathbf{x}} \in \mathbb{R}^{d+1} \ ,$$

with $d = 1, 2$ and the local wave speed

$$c^{-2}(z, \mathbf{x}) = c_0^{-2}\left[1 + \sigma_0 \mu\left(\frac{z}{l_z}, \frac{\mathbf{x}}{l_x}\right)\right].$$

Here z and $\mathbf{x} \in \mathbb{R}^d$ are, respectively, the coordinates along and transverse to the direction of propagation, and $\vec{\mathbf{x}} = (z, \mathbf{x})$. The random function μ models the fluctuations in the propagation speed.

When the characteristic scale of variation in the transverse direction l_x is large compared to l_z then we have a layered random medium. When $l_x = l_z = l$ then we have essentially isotropic randomness.

The three regimes of random wave propagation

- Layered: Very strong scattering in direction of propagation. Wave localization, long wave codas

- Wave transport: Wave energy "diffuses" by radiative transport. The transport mean free path

- The paraxial or parabolic regime: one-way wave propagation for beams, with scattering into lateral directions and no backscattering

- Layered and paraxial are approximations that have very well developed mathematical theories. Real world phenomena are somewhere in between

Paraxial or parabolic approximation

We consider wave fields propagating mainly in the z direction

$$u(t, \mathbf{x}, z) = \frac{1}{2\pi} \int e^{i\omega(z/c_0 - t)} \psi(z, \mathbf{x}; \omega/c_0) d\omega$$

The complex amplitude $\psi(z, \mathbf{x}; k)$ satisfies the Helmholtz equation

$$2ik\psi_z + \Delta_{\mathbf{x}}\psi + k^2(n^2 - 1)\psi = -\psi_{zz}.$$

Here $k = \omega/c_0$ is the wavenumber and $n(\mathbf{x}, z) = c_0/c(\mathbf{x}, z)$ is the random index of refraction relative to a reference speed c_0. The fluctuations of the refraction index have the form

$$n^2(\mathbf{x}, z) - 1 = \sigma_0 \mu \left(\frac{z}{l}, \frac{\mathbf{x}}{l} \right)$$

They are a stationary random field with mean zero, variance σ_0^2 and correlation length l. The normalized and dimensionless covariance is given by

$$R(z, \mathbf{x}) = E\{\mu(z + z', \mathbf{x} + \mathbf{x}')\mu(z', \mathbf{x}')\}.$$

On the numerical simulation of μ

1. Write (in 1D): $R(x) = (1/2\pi)\int dk e^{ikx}\widehat{R}(k)$, with $\widehat{R}(k)$ the power spectral density, and discretize the integral with mesh size Δk

2. Generate independent identically distributed complex random variables $\widehat{\mu}_n$ with mean zero and variance $\widehat{R}(n\Delta k)\Delta k/2\pi$, and so that $\overline{\widehat{\mu}_n} = \widehat{\mu}_{-n}$

3. The process $\mu_{\Delta k}(x) = \sum_n e^{in\Delta k x}\widehat{\mu}_n$ is an approximate realization of $\mu(x)$

Scales

- L_z, the characteristic distance in the direction of propagation.

- $L_{\mathbf{x}}$, the length scale in the directions transverse to the direction of propagation. This is typically taken to be the width of the propagating beam.
- $k_0 = 2\pi/\lambda_0$, the central wavenumber corresponding to the central wavelength λ_0.
- l, the correlation length of the random medium. It characterizes the dominant spatial scale of the random fluctuations.

- σ_0, the dimensionless standard deviation of the random fluctuations in the medium.

In the asymptotic regimes that we consider here L_z and L_x are large compared to l and λ_0, and σ_0 is small.

Scaled, dimensionless wave equation

We obtain the dimensionless form of the equation by introducing dimensionless variables by $\mathbf{x} = L_{\mathbf{x}}\mathbf{x}'$, $z = L_z z'$, $k = k_0 k'$ and rewriting it as

$$2ik\frac{\partial \psi}{\partial z} + \frac{L_z}{k_0 L_{\mathbf{x}}^2}\Delta_{\mathbf{x}}\psi + k^2 k_0 L_z \sigma_0 \mu\left(\frac{zL_z}{l}, \frac{\mathbf{x}L_{\mathbf{x}}}{l}\right)\psi = -\frac{1}{L_z k_0}\frac{\partial^2 \psi}{\partial z^2},$$

after dropping the primes. We identify now the following three, usually small, dimensionless parameters in the problem:

- $\varepsilon = \dfrac{l}{L_z}$, the ratio of the correlation length to the propagation distance,
- $\delta = \dfrac{l}{L_{\mathbf{x}}}$, the ratio of the correlation length to the transverse length scale, which is usually the beam width,
- $\theta = \dfrac{L_z}{k_0 L_{\mathbf{x}}^2} = \dfrac{\lambda_0 L_z}{2\pi L_{\mathbf{x}}^2}$, the reciprocal of the Fresnel number, the ratio of the diffraction focal spot of the beam to its width.

In terms of these parameters we have

$$2ik\psi_z + \theta\Delta_{\mathbf{x}}\psi + \frac{k^2\sigma_0\delta^2}{\theta\varepsilon^2}\mu(\frac{z}{\varepsilon}, \frac{\mathbf{x}}{\delta})\psi = -\frac{\theta\varepsilon^2}{\delta^2}\psi_{zz}.$$

We assume that ε is the smallest parameter in the problem. It then follows formally, but it is quite difficult to prove, that the ψ_{zz} term is a lower order term and can be neglected.

$$2ik\psi_z + \theta\Delta_{\mathbf{x}}\psi + \frac{k^2\sigma\delta}{\theta\sqrt{\varepsilon}}\mu\left(\frac{z}{\varepsilon}, \frac{\mathbf{x}}{\delta}\right)\psi = 0\,, z > 0$$

with ψ at $z = 0$ given and where

$$\sigma = \frac{\sigma_0\delta}{\varepsilon^{3/2}}.$$

This scaled noise strength parameter is assumed to be independent of ε and δ as these parameters tend to zero in the asymptotic analysis.

The white noise limit

We consider the limit $\varepsilon \to 0$ while δ and θ are fixed. This means that ε is the smallest of the three parameters $\varepsilon, \theta, \delta$. Assume that the CLT applies to the random field μ:

$$\lim_{\varepsilon \to 0} \frac{1}{\sqrt{\varepsilon}} \int_0^z \mu\left(\frac{s}{\varepsilon}, \mathbf{x}\right) ds = B(z, \mathbf{x}),$$

weakly in law, where B is a Brownian random field parameterized by $\mathbf{x}$. This means that for any test function $h(\mathbf{x})$, in law

$$\frac{1}{\sqrt{\varepsilon}} \int_0^z \mu_h(s/\varepsilon) ds \mapsto B_h(z), z \geq 0,$$

$$\mu_h(z) = \int_{\mathbb{R}^d} \mu(z, \mathbf{x}) h(\mathbf{x}) d\mathbf{x} \ , \quad B_h(z) = \int_{\mathbb{R}^d} B(z, \mathbf{x}) h(\mathbf{x}) d\mathbf{x}.$$

The random field $B(z, \mathbf{x})$ is Gaussian with mean zero and

$$E\{B(z_1, \mathbf{x}_1) B(z_2, \mathbf{x}_2)\} = R_0(|\mathbf{x}_1 - \mathbf{x}_2|) \min\{z_1, z_2\}.$$

Here R_0 is the integrated correlation function $R_0(\mathbf{x}) = \int_{-\infty}^{\infty} R(z, \mathbf{x}) dz$.

The Ito-Schrödinger equation

In the white noise limit $\varepsilon \to 0$ the solution of the random partial differential equation converges in law to the process defined by the stochastic partial differential equation

$$2ikd_z\psi + \theta\Delta_{\mathbf{x}}\psi dz + \frac{k^2\sigma\delta}{\theta}\psi \circ d_z B\left(\frac{\mathbf{x}}{\delta}, z\right) = 0$$

given here in the Stratonovich form. The Itô form is

$$2ikd_z\psi + \theta\Delta_{\mathbf{x}}\psi dz + \frac{ik^3\sigma^2\delta^2}{4\theta^2}R_0(0)\psi dz + \frac{k^2\sigma\delta}{\theta}\psi d_z B\left(\frac{\mathbf{x}}{\delta}, z\right) = 0.$$

There are two small parameters left in the Itô-Schrödinger equation after we have taken the white-noise limit – the reciprocal Fresnel number θ and the non-dimensional correlation length δ.

We can consider the following limits.

High and low frequency; lateral diversity

- The low frequency limit and large lateral diversity limit: $\delta \to 0$ with θ fixed,

- the high frequency or geometric asymptotics limit followed by the large lateral diversity limit: $\theta \ll \delta \ll 1$, that is, $\theta \to 0$ followed by $\delta \to 0$, and

- the combined scaling limit: $\theta \sim \delta \ll 1$ with $\theta \to 0$ and $\delta \to 0$ simultaneously.

We refer to the limit $\theta \to 0$ as the high frequency limit and to the limit $\delta \to 0$ as the limit of large lateral diversity.

We see that if we pass to the limit $\delta \to 0$ with a fixed $\theta > 0$ we arrive at the homogeneous Schrödinger equation

$$2ik\psi_z + \theta\Delta_{\mathbf{x}}\psi = 0.$$

This is because we have an a priori bound $\|\psi(t)\|_{L^2} = \|\psi_0\|_{L^2}$ and for any deterministic test function $\eta(z,\mathbf{x})$ we have by the Itô isometry

$$E\left[\frac{k^2\sigma\delta}{\theta}\int_0^z\int \eta(s,\mathbf{x})\psi(s,\mathbf{x})d_zB\left(\frac{\mathbf{x}}{\delta},s\right)d\mathbf{x}\right]^2$$
$$=\left(\frac{k^2\sigma\delta}{\theta}\right)^2 E\int_0^z\int \eta(s,\mathbf{x})\eta(s,\mathbf{x}')\psi(s,\mathbf{x})\psi(s,\mathbf{x}')R_0\left(\frac{\mathbf{x}-\mathbf{x}'}{\delta}\right)d\mathbf{x}d\mathbf{x}'ds \to 0 \text{ as } \delta \to 0.$$

A similar bound holds for the third term and therefore convergence in probability follows.

Phase space

In the high frequency limit $\theta \to 0$ (whether coupled with the limit $\delta \to 0$, or not) solutions of the Itô-Schrödinger equation become oscillatory in time and space. Therefore, rather than studying the limit of the solution itself we consider the limits of its Wigner transform which resolves the wave energy of oscillatory fields in the phase space and (unlike the spatial energy density) satisfies a closed evolution equation.

We define the spatial Fourier transform and its inverse by

$$\widehat{f}(\mathbf{k}) = \int d\mathbf{x} e^{-i\mathbf{k}\cdot\mathbf{x}} f(\mathbf{x}) \;, \quad f(\mathbf{x}) = \int \frac{d\mathbf{k}}{(2\pi)^d} e^{i\mathbf{k}\cdot\mathbf{x}} \widehat{f}(\mathbf{k}) \,,$$

where $d = 1$ or 2 is the number of transverse spatial dimensions.

The Wigner transform relative to the scale θ is

$$W_\theta(z, \mathbf{x}, \mathbf{p}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\mathbf{p}\cdot\mathbf{y}} \psi(\mathbf{x} - \frac{\theta\mathbf{y}}{2}, z) \overline{\psi(\mathbf{x} + \frac{\theta\mathbf{y}}{2}, z)} d\mathbf{y}$$

The Wigner distribution is real, may be interpreted as phase space wave energy. It is well suited for random media.

Stochastic transport equation

Using the Itô calculus we find from the Ito-Schrödinger equation that the scaled Wigner distribution satisfies the stochastic transport equation

$$dW_\theta(z,\mathbf{x},\mathbf{p}) + \frac{\mathbf{p}}{k}\cdot\nabla_{\mathbf{x}} W_\theta(z,\mathbf{x},\mathbf{p})dz = \frac{k^2\sigma^2\delta^2}{4\theta^2}\int \left(W_\theta\left(z,\mathbf{x},\mathbf{p}+\frac{\theta\mathbf{q}}{\delta}\right) - W_\theta(z,\mathbf{x},\mathbf{p})\right) \frac{\widehat{R}_0(\mathbf{q})d}{(2\pi)^d}$$

$$+\frac{ik\sigma\delta}{2\theta}\int \frac{d\mathbf{q}}{(2\pi)^d} e^{i\mathbf{q}\cdot\mathbf{x}/\delta}\left(W_\theta\left(z,\mathbf{x},\mathbf{p}-\frac{\theta\mathbf{q}}{2\delta}\right) - W_\theta\left(z,\mathbf{x},\mathbf{p}+\frac{\theta\mathbf{q}}{2\delta}\right)\right) d\widehat{B}(\mathbf{q},z).$$

We do the high frequency and large diversity limits with the Itô-Wigner equation as a starting point.

We note that the L^2 norm of the Wigner distribution is conserved

$$\|W_\theta(z)\|_{L^2(\mathbb{R}^{2d})} = \|W_\theta(0)\|_{L^2(\mathbb{R}^{2d})}$$

High frequency limit

When we take the high frequency limit we find that W_θ converges weakly to W_δ satisfying the Itô-Liouville equation

$$dW_\delta(z,\mathbf{x},\mathbf{p})+\frac{\mathbf{p}}{k}\cdot\nabla_\mathbf{x}W_\delta(z,\mathbf{x},\mathbf{p})dz+\frac{k^2\sigma^2}{8}R_0''(0)\triangle_\mathbf{p}W_\delta dz=-\frac{k\sigma}{2}d\nabla_\mathbf{x}B\left(\frac{\mathbf{x}}{\delta},z\right)\cdot\nabla_\mathbf{p}W_\delta.$$

We remark that $R''(0)<0$ so that this equation is well-posed.

This SPDE is connected to stochastic flows where solutions of SDE's play the role of characteristics (Kunita).

The limiting Wigner distribution solves a stochastic PDE, in which the coefficient of the random term fluctuates on the small scale δ. When we subsequently take the limit of large lateral diversity we find that the limiting Wigner distribution actually becomes deterministic. We refer to this as the stabilization of the Wigner distribution. Define W as the deterministic solution of

$$\frac{\partial W}{\partial z}(z,\mathbf{x},\mathbf{p}) + \frac{\mathbf{p}}{k}\cdot\nabla_{\mathbf{x}} W(z,\mathbf{x},\mathbf{p}) + \frac{k^2\sigma^2}{8} R_0''(0)\triangle_{\mathbf{p}} W = 0.$$

Moment formula

There are two simple and practical items to remember when with waves in random media and how affect TR and imaging calculations.

One is the moment formula:

$$E\{\overline{\widehat{G}(\mathbf{x}_r,\mathbf{y},\omega)}\widehat{G}(\mathbf{x}_r,\mathbf{y}^S,\omega)\} \approx \overline{\widehat{G}_0(\mathbf{x}_r,\mathbf{y},\omega)}\widehat{G}_0(\mathbf{x}_r,\mathbf{y}^S,\omega)e^{\frac{-k^2\xi^2 a_e}{2L^2}}$$

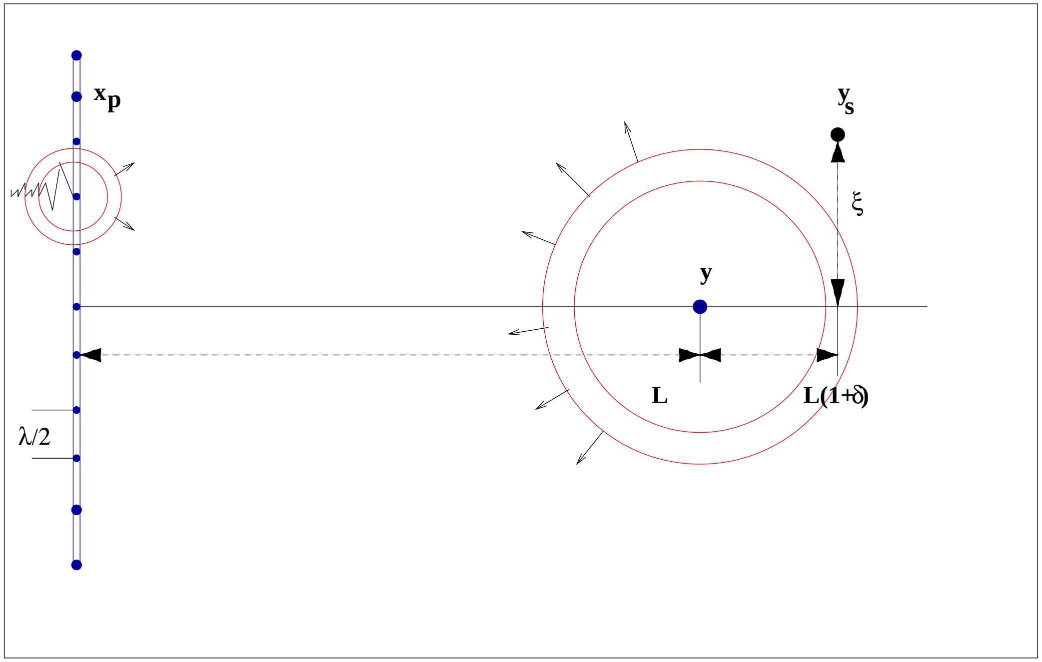
The other is statistical stability: When integrating over a sufficiently wide frequency band we have

$$\int d\omega\overline{\widehat{G}(\mathbf{x}_r,\mathbf{y},\omega)}\widehat{G}(\mathbf{x}_r,\mathbf{y}^S,\omega) \approx \int d\omega E\{\overline{\widehat{G}(\mathbf{x}_r,\mathbf{y},\omega)}\widehat{G}(\mathbf{x}_r,\mathbf{y}^S,\omega)\}$$

Thus, we under favorable conditions we have for example

$$\int d\omega \sum_{\mathbf{x}_r}\overline{\widehat{G}(\mathbf{x}_r,\mathbf{y},\omega)}\widehat{G}(\mathbf{x}_r,\mathbf{y}^S,\omega) \approx \int d\omega \sum_{\mathbf{x}_r}\overline{\widehat{G}_0(\mathbf{x}_r,\mathbf{y},\omega)}\widehat{G}_0(\mathbf{x}_r,\mathbf{y}^S,\omega)e^{\frac{-k^2\xi^2 a_e}{2L^2}}$$

Time Reversal Schematic



Range: L, Carrier wavelength λ, Array size $a = (N-1)\lambda/2$.
Source at y, Search point at y_s, Transducers at x_p.
Remote sensing regime: $\lambda << a << L$.
Random medium: Correlation length $l << L$, fluctuation strength $\sigma << 1$.

Remarks on TR in RM

- Resolution in time reversal: $\frac{\lambda L}{a}$, cross-range. It is the same as the Rayleigh resolution of optical instruments

- Super-resolution in random media because of multiple scattering: $\lambda L/a_e$, cross-range. The effective aperture a_e can be much larger that the physical aperture a. In random media, resolution is better than the diffraction limit

- Statistical stability (self-averaging) of time-reversed and back-propagated field. Broad-band and narrow-band signals. Super-resolution is observed only in regimes where there is statistical stability

On the plane of the source, at a point with transverse coordinates ξ, the time time harmonic field is

$$\psi^B(L,\xi,k) = \int G_\theta(L,x,\xi;k)\overline{G_\theta(L,\eta,x;k)\psi_0(\eta,k)}\chi_A(x)dxd\eta$$

where G_θ is the (random) Green's function. In the time domain it is

$$\Psi^B(L,\xi,t) = \int e^{-i\omega t}\psi^B(L,\xi,\frac{\omega}{c_0})d\omega$$

Because of the form of this field, and for many other reasons, we introduce and use the Wigner distribution of ψ

$$W_\theta(z,x,p) = \int \frac{dy}{(2\pi)^2}e^{ip\cdot y}\psi(z,x-\frac{\theta y}{2},k)\overline{\psi(z,x+\frac{\theta y}{2},k)}$$

and note that ψ^B can be written entirely in terms of W_θ.

High frequency limit $\theta \to 0$

The Wigner distribution satisfies a linear stochastic equation, the Ito-Wigner equation, that comes from the Ito-Schrödinger equation using the Ito calculus. In the high frequency limit the Wigner process converges weakly to the solution of the Ito-Liouville equation

$$d_z W + (\frac{p}{k} \cdot \nabla_x W - \frac{k^2 D}{2} \Delta_p W) dz + \frac{k}{2} \nabla_p W \cdot \nabla_x d_z B(\frac{x}{\delta}, z) = 0$$

where $D = -R_0''(0)/4$ and the wave number scales out: $W = W(z, x, p/k; k = 1)$. The expected value $E\{W\}$ solves the PDE

$$W_z + \frac{p}{k} \cdot \nabla_x W - \frac{k^2 D}{2} \Delta_p W = 0$$

with given initial conditions $W(0, x, p; k)$.

The process W depends on δ but $E\{W\}$ does not.

If we take a source field that is a directed beam

$$e^{ip_0 \cdot x/\theta} \psi_0(\frac{x}{\sigma_s}, k),$$

with σ_s the lateral extent of the source, then in the white-noise ($\epsilon \to 0$) and high-frequency ($\theta \to 0$) limits we have

$$E\{\psi^B(L, \xi, k)\} = \psi_0(\cdot, -k) * \mathcal{W}(\cdot)(\xi)$$

where $\mathcal{W}$ is the point spread function

$$\mathcal{W}(\eta) = \left(\frac{k}{2\pi L}\right)^2 \hat{\chi}_A(\frac{\eta k}{L}) e^{-\eta^2/(2\sigma_M^2)}$$

and

$$\sigma_M = \frac{L}{k a_e}\ , \quad a_e = \sqrt{\frac{DL^3}{3}}$$

Here $a_e = a_e(L)$ is the effective aperture of the array.

Interpretation of the point spread function

If there is no scattering medium then $D = 0$ and

$$\mathcal{W}(\eta) = \left(\frac{k}{2\pi L}\right)^2 \widehat{\chi}_A(\frac{\eta k}{L})$$

For a square aperture $A = [-\frac{a}{2}, \frac{a}{2}]^2$

$$\mathcal{W}(\eta) = \mathcal{W}(\eta_1, \eta_2) = \frac{1}{\pi^2 \eta_1 \eta_2} \sin(\frac{\eta_1 ka}{2L}) \sin(\frac{\eta_2 ka}{2L})$$

The first zero of the sine function is at

$$\eta_F = \frac{2\pi L}{ka} = \frac{\lambda L}{a} = \text{Rayleigh resolution}$$

If we define $\sigma_F = L/ka$, the Fresnel spot size, then when $\sigma_F << \sigma_M$, or $a >> a_e$, multipathing does not alter the refocused spot size of diffraction theory.

But if $a_e >> a$ then the point spread function is

$$\mathcal{W} \approx \left(\frac{a}{\sqrt{2\pi} a_e}\right)^2 \frac{e^{-\eta^2/(2\sigma_M^2)}}{2\pi\sigma_M^2}$$

Self-averaging

When is the time-reversed, back-propagated field self-averaging? This is a fundumental issue because it determines when super-resolution is observable.

In the present setting there are two results:

- If the source is localized, $\sigma_s \sim \theta$, then, in the limit $\delta \to 0$, the time harmonic field ψ^B is self-averaging

$$\lim_{\delta \to 0} E\{(\psi^B - E\{\psi^B\})^2\} = 0$$

- If the source is distributed, $\sigma_s >> \theta$, then only in the time domain, that is for $\Psi^B(L, \xi, t)$, we have self-averaging in mean square sense as $\delta \to 0$.

- What does $\delta \to 0$ mean? Provides cross-range diversity in multipathing.

The self-averaging is based on the following theorem for the Ito-Liouville process (with $k = 1$) defined by

$$d_z W + (p \cdot \nabla_x W - \frac{D}{2}\Delta_p W)dz + \frac{1}{2}\nabla_p W \cdot \nabla_x d_z B(\frac{x}{\delta}, z) = 0$$

with $W(0, x, p) = \chi_A(x)$:

For any $z > 0$ the integral

$$J_\delta(z, x) = \int W_\delta(z, x, p)dp$$

exists and

$$\lim_{\delta \to 0} E\{(J_\delta - E\{J_\delta\})^2\} = 0$$

where $E\{J_\delta\}$ is independent of δ.

This is proved by using properties of the SDE's (random characteristics) through which the Ito-Liouville equation can be solved.

Time-reversed, back-propagated pulse

In the time domain and for a distributed source, the self-averaging field, in the white-noise and high-frequency limit, is given by

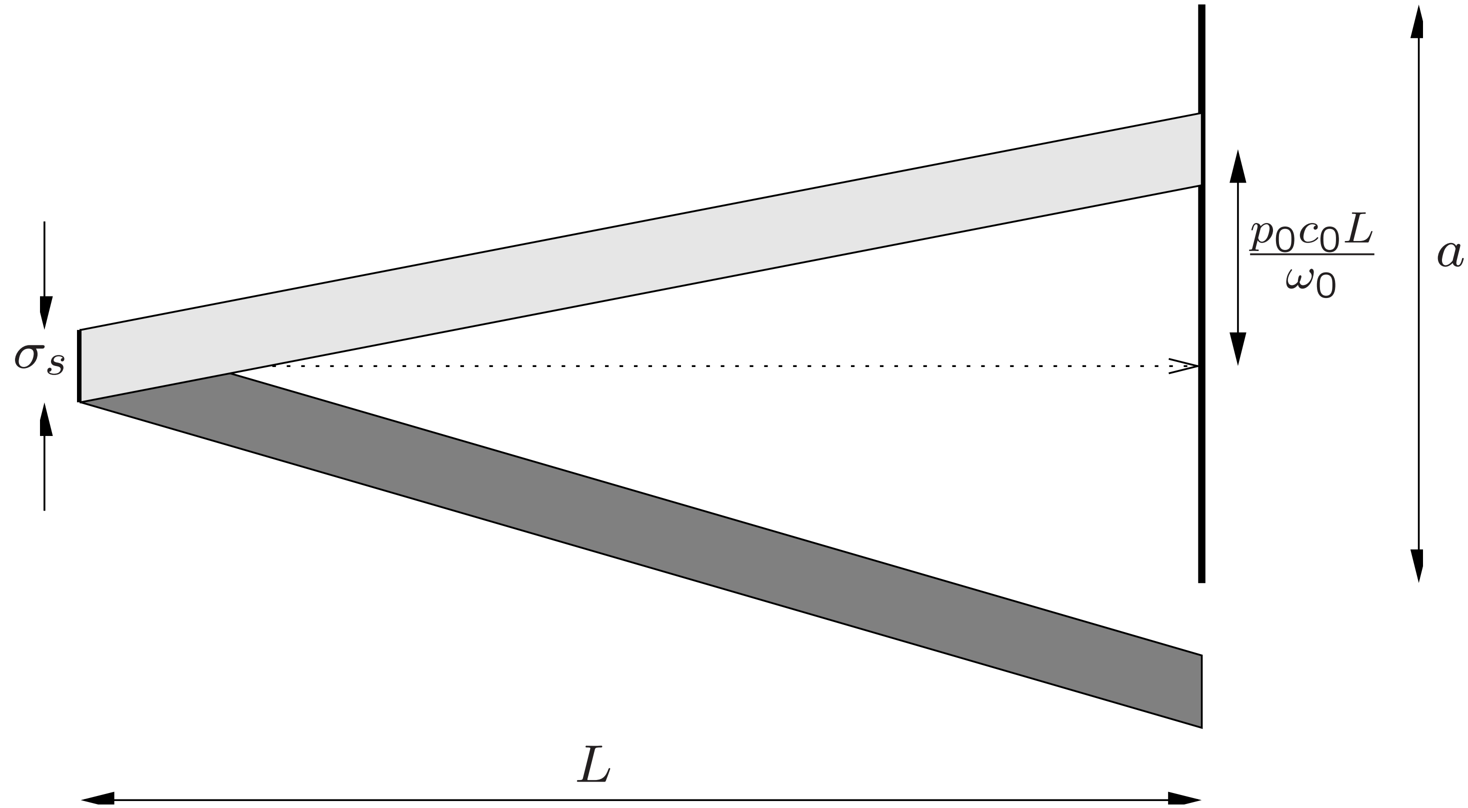
$$\Psi^B(L,\xi,t) = e^{-i(p_0\cdot\xi+\omega_o t)}\psi_0(\xi) \cdot \int_{\{|\omega|<\Omega\}} \frac{d\omega}{2\pi} e^{-i\omega t}\widehat{g}(-\omega)\ \chi_A * \left(\frac{e^{-x^2/2a_e^2}}{2\pi a_e^2}\right)\left(\frac{Lc_0p_0}{\omega_0+\omega}\right)$$

When $a_e << a$, that is, no multipathing, then

$$\Psi^B(L,\xi,t) \sim e^{-i(p_0\cdot\xi+\omega_o t)}\psi_0(\xi) \cdot \int_{\{|\omega|<\Omega\}} \frac{d\omega}{2\pi} e^{-i\omega t}\widehat{g}(-\omega)\ \chi_A \left(\frac{Lc_0p_0}{\omega_0+\omega}\right)$$

In this case, if the beam lands entirely withing the TRM then the time-reversed and back-propagated pulse is

$$e^{-i(p_0\cdot\xi+\omega_o t)}\psi_0(\xi)g(-t)$$



A directed field propagates from a distributed source of size σ_s toward the time reversal mirror of size a. The time-reversed, back-propagated field depends on the location of the mirror relative to the direction of the propagating beam.

Time-reversed, back-propagated pulse with multipathing

When multipathing is strong, $a_e >> a$, then the self-averaging time-reversed and back-propagated pulse is given by

$$\Psi^B(L,\xi,t) \sim e^{-i(p_0\cdot\xi+\omega_o t)}\psi_0(\xi) \cdot \left(\frac{a}{\sqrt{2\pi}a_e}\right)^2 \int_{\{|\omega|<\Omega\}} \frac{d\omega}{2\pi} e^{-i\omega t}\widehat{g}(-\omega)\ e^{-\frac{1}{2}\left(\frac{Lc_0p_0}{a_e(\omega_0+\omega)}\right)^2}$$

Note that, remarkably, this expression **is almost independent of the time reversal mirror**!

Use this formula to estimate the most important quantity in time reversal with strong multipathing: the effective aperture a_e.

Point the beam in different directions toward the TRM, measure the time reversed pulse and estimate a_e by fitting to the formula.

Summary and conclusions

- Time reversal in a random medium is important because of super-resolution and self-averaging, which are phenomena that are difficult to analyze and understand quantitatively, and require interesting mathematics.

- Applications abound, are very exciting and limited only by the **hardware**, our imagination, and also our analytical understanding: Direct TR applications, Imaging, Communications.

Imaging in Random Media

George Papanicolaou
Department of Mathematics, Stanford University
http://math.stanford.edu/

$Id: NSF-CBMS, CAAM Rice, May 12-16, 2008 $

Coherent interferometry for imaging in random media

Kirchhoff or travel time migration

Assume that the background velocity is known. Denote the deterministic background Green's function by $\widehat{G}_0(\mathbf{x},\mathbf{y},\omega) = \frac{e^{i\omega\tau(\mathbf{x},\mathbf{y})}}{4\pi|\mathbf{x}-\mathbf{y}|}$. We can then use the following imaging functional for the reflectivity $\rho(\mathbf{y}^S)$:

$$I^{KM}(\mathbf{y}^S) = \sum_{\mathbf{x}_s,\mathbf{x}_r} P(\mathbf{x}_r,\mathbf{x}_s,\tau(\mathbf{x}_s,\mathbf{y}^S)+\tau(\mathbf{y}^S,\mathbf{x}_r))$$

Here $\tau(\mathbf{x},\mathbf{y}) = |\mathbf{x}-\mathbf{y}|/c_0$ is the travel time from $\mathbf{x}$ to $\mathbf{y}$ when the speed of propagation is c_0.

This does not work in clutter because the deterministic travel time cannot deal with the delay spread in the traces. The delay spread is due to the scattering from the random inhomogeneities.

Delay spread manifests itself in the frequency domain as random phases. To avoid this random phase problems in Kirchhoff migration imaging we mimic physical time reversal by computing cross-correlations of data traces, the interferograms, and summing

$$I^{INT}(\mathbf{y}^S) = \sum_{\mathbf{x}_r,\mathbf{x}_{r'}} P(\mathbf{x}_r,\cdot) *_t P(\mathbf{x}_{r'},-\cdot)\big|_{\tau(\mathbf{x}_r,\mathbf{y}^S)-\tau(\mathbf{x}_{r'},\mathbf{y}^S)}$$

The interferograms are given by

$$P(\mathbf{x}_r,\cdot) *_t P(\mathbf{x}_{r'},-\cdot)(t) = \int_{-\infty}^{\infty} P(\mathbf{x}_r,s)P(\mathbf{x}_{r'},s-t)ds$$

In the frequency domain we have

$$I^{INT}(\mathbf{y}^S) = \int d\omega \left|\sum_{\mathbf{x}_r} \widehat{P}(\mathbf{x}_r,\omega)e^{-i\omega\tau(\mathbf{x}_r,\mathbf{y}^S)}\right|^2$$

In the frequency domain we have

$$I^{INT}(\mathbf{y}^S) = \int d\omega \left| \sum_{\mathbf{x}_r} \widehat{P}(\mathbf{x}_r, \omega) e^{-i\omega\tau(\mathbf{x}_r, \mathbf{y}^S)} \right|^2$$

This is almost Matched Field Imaging

$$I^{MF}(\mathbf{y}^S) = \int d\omega \left| \sum_{\mathbf{x}_r} \overline{\widehat{P}(\mathbf{x}_r, \omega)} \widehat{G}_0(\mathbf{x}_r, \mathbf{y}^S, \omega) \right|^2 , \quad \widehat{G}_0(\mathbf{x}, \mathbf{y}, \omega) = \frac{e^{i\omega\tau(\mathbf{x},\mathbf{y})}}{4\pi|\mathbf{x} - \mathbf{y}|}$$

that is widely used in sonar and elsewhere in more general situations (waveguides, enclosures, etc) with a suitable $\widehat{G}_0$.

Decoherence distance X_d and decoherence frequency Ω_d

The trace cross-correlation

$$P(\mathbf{x}_r,\cdot) *_t P(\mathbf{x}_{r'},-\cdot)(t)$$

does not have a peak if $|\mathbf{x}_r - \mathbf{x}_{r'}| > X_d$.

The phases of $\widehat{P}(\mathbf{x}_r,\omega_1)$ and $\widehat{P}(\mathbf{x}_r,\omega_2)$ decorrelate when $|\omega_1 - \omega_2| > \Omega_d$.

Both X_d and Ω_d can be ESTIMATED from the array data directly.

Use the Coherent Interferometric imaging functional:

$$I^{CINT}(\mathbf{y}^S; X_d, \Omega_d) = \int\int_{|\omega_1-\omega_2|\le\Omega_d} d\omega_1 d\omega_2 \sum\sum_{|\mathbf{x}_r-\mathbf{x}'_r|\le X_d} \widehat{P}(\mathbf{x}_r,\omega_1)\overline{\widehat{P}(\mathbf{x}'_r,\omega_2)}e^{-i(\omega_1\tau(\mathbf{x}_r,\mathbf{y}^S)-\omega_2\tau(\mathbf{x}'_r,\mathbf{y}^S))}$$

If we take $X_d = a$ and $\Omega_d = B$, which means that there is no smoothing, then the CINT functional is just the Kirchhoff migration functional squared: $I^{CINT} = (I^{KM})^2$. The case $\Omega_d = 0$, suitably interpreted, is incoherent interferometry.

Adaptive Selection of X_d and Ω_d

For CINT to be effective we need to be able to determine the decoherence length and frequency adaptively. We do this by minimizing the bounded variation norm of the (random) functional I^{CINT}

$$\{X_d^*, \Omega_d^*\} = \text{argmin}_{X_d, \Omega_d} ||I^{CINT}(\cdot; X_d, \Omega_d)||_{BV}$$

where

$$||f||_{BV} = \int |f(\mathbf{y})| d\mathbf{y} + \alpha \int |\nabla f(\mathbf{y})| d\mathbf{y}$$

The bounded variation norm is used because it is smoothing on small scales but is respectful of large scale features (discontinuities). The smoothing is limited by the L^1 norm. Other sparsity-type norms can be used, such as entropy norms.

What about denoising the data first and then migrating?

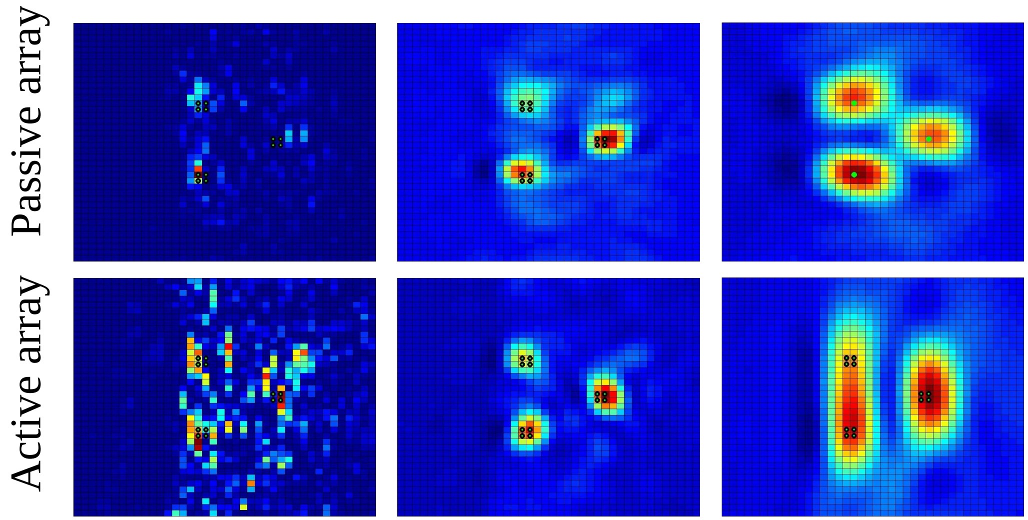
The array data P, not the image, could be denoised by minimizing

$$\alpha||P - Q||_{PROX} + ||Q||_{REG}$$

over Q. This is an expensive calculation for large array data sets.

The denoising can also be done by harmonic analysis methods: decompose the data P in some well chosen basis (ridgelets?), threshold the Fourier coefficients below some level and reconstruct to get the denoised data Q.

After the denoising one can do Kirchhoff migration with Q as array data.



Coherent Interferometry images in random media with $s = 3\%$.

Left Figures: $X_d = a$, $\Omega_d = B$ (Kirchhoff Migration, no smoothing)

Middle Figures: $X_d = X_d^*$, $\Omega_d = \Omega_d^*$ (Adaptively selected optimal smoothing)

Right Figures: $X_d < X_d^*$, $\Omega_d < \Omega_d^*$ (Too much smoothing)

Comments on the CINT results

- Without smoothing there is no statistical stability of the image: Different realizations of the random medium give different images. Smoothing, especially in frequency, gives stable but blurred images.

- Statistical stability of the image is very important because it allows further processing with deblurring methods. We have used Level Set Deblurring methods successfully, provided that we have a good estimate of the amount of blurring.

- The optimal decoherence frequency Ω_d^* is not known and it is determined adaptively, as explained above. So is the decoherence distance X_d^*.

Resolution theory for specific models

A resolution theory can be developed based on several assumptions about the random medium and the propagation regime.

Such assumptions are NOT used in the numerical simulations.

- With the paraxial approximation, the white noise limit, and a high frequency expansion we reduce all theoretical calculations to the use of one relatively simple formula obtained from the random Schrödinger equation: a second order moment formula.

- One other regime where analytical results can be obtained: Layered media. In no other regime do we have, or expect, analytical results.

Summary: Resolution limits

	Determin	Known Rand (TR)	Unknown Rand (CINT)
Range	$\frac{c_0}{B}$	$\frac{c_0}{B}$	$\frac{c_0}{\Omega_d}$
Cross Range	$\frac{c_0 L}{\omega a}$	$\frac{c_0 L}{\omega a_e} \sim X_d$	$\frac{c_0 L}{\omega X_d} \sim a_e$

Resolution limits in deterministic media and in random media, when the random medium is known as in physical time reversal, and when the random medium is not known as in coherent interferometry.

Role of coherence in array imaging

Coherence is essential in array imaging. A more physical, and more conventional, way to measure coherence is through the transport mean free path $l^\star$. In the regime where waves energy propagates by diffusion, the diffusion coefficient is given by

$$D = \frac{c_0 l^\star}{3}.$$

If the transport mean free path $l^\star$ is small compared to the range L of the object to be imaged then migration methods, including CINT, will not work. In our numerical simulations $l^\star$ is of the order of L, which is the regime where we expect CINT to be effective.

Optimal illumination in array imaging is up to now aimed at DETECTION. That is, if $\widehat{f}(\mathbf{x}_s, \omega)$ is the signal in the frequency domain that is emitted at $\mathbf{x}_s$ then the signal received at $\mathbf{x}_r$ is given by

$$\sum_{s=1}^{N_s} \widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega) \widehat{f}(\mathbf{x}_s, \omega).$$

The total power received at the array is given by

$$\mathcal{P} = \int_{|\omega - \omega_0| \leq B} d\omega \sum_{r=1}^{N_r} \left| \sum_{s=1}^{N_s} \widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega) \widehat{f}(\mathbf{x}_s, \omega) \right|^2 .$$

We want to maximize this functional over all illumination signals $\widehat{f}(\mathbf{x}_s, \omega)$ with the normalization

$$\int_{|\omega - \omega_0| \leq B} d\omega \sum_{s=1}^{N_s} \left| \widehat{f}(\mathbf{x}_s, \omega) \right|^2 = 1$$

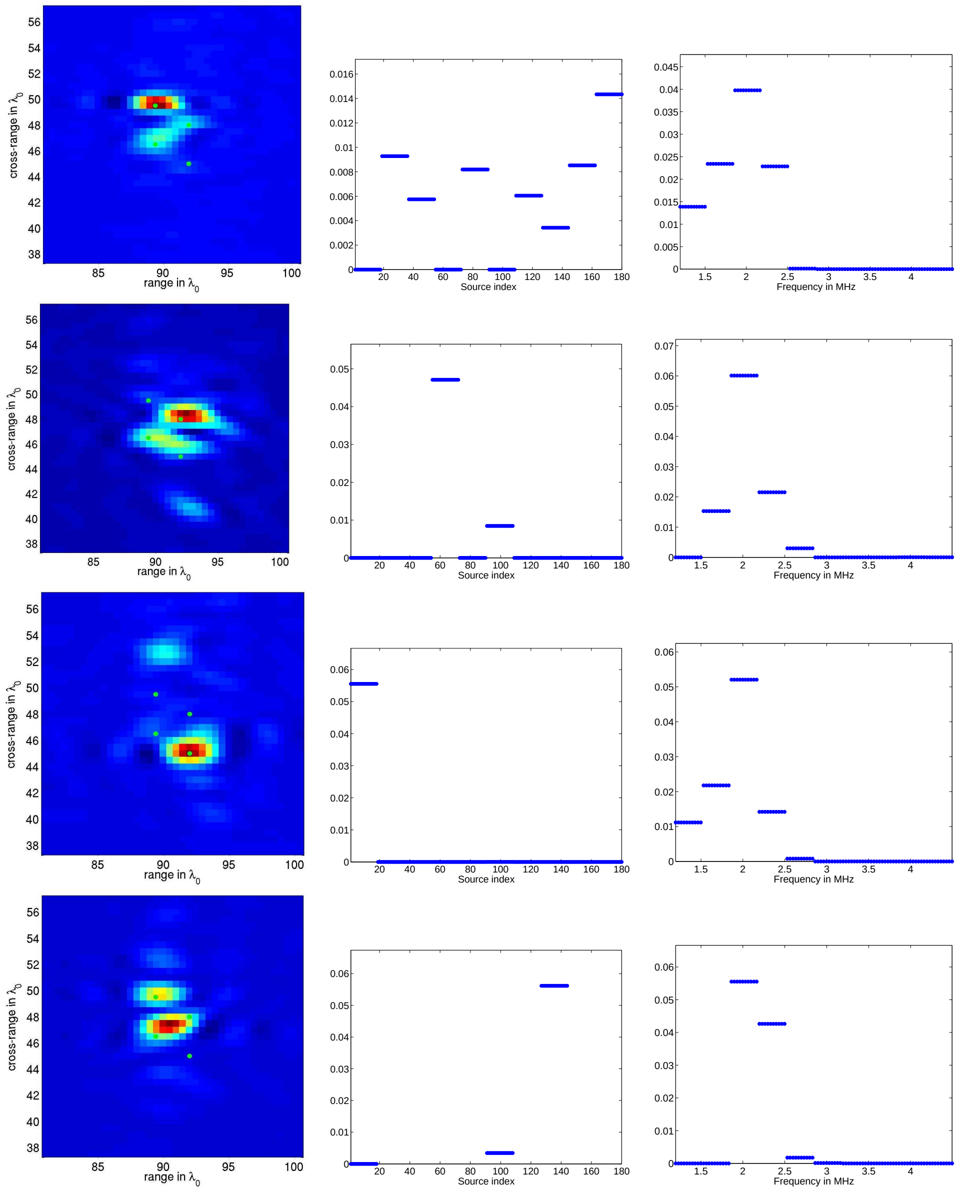
Optimal illumination and waveform design

In general, Optimal Illumination for Detection (SVD) produces images with bad resolution.

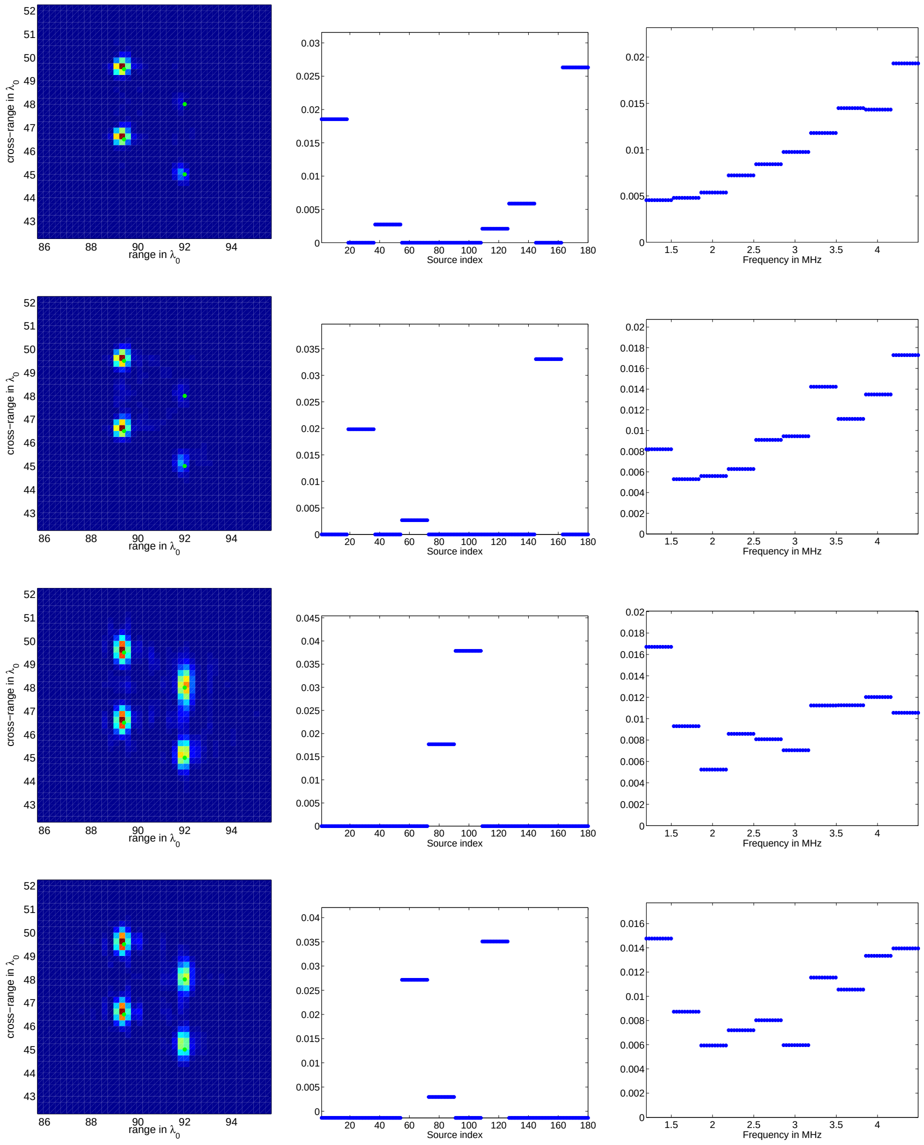
Just like in Adaptive Coherent Interferometry we can, however, introduce an optimal illumination objective that is tied to the resolution of the image itself.

We now show the results of numerical simulations using this approach to optimal illumination.

Optimal Illumination, Random Medium



Left: image; Center, weights; Right: pulse.



Left: image; Center: weights; Right: pulse

We compute first the weighted average of the migrated traces

$$\widehat{m}(\mathbf{x}_r, \mathbf{y}^S, \omega) = \sum_{s=1}^{N_s} w_s \widehat{f}_B(\omega - \omega_0) \widehat{\Pi}(\mathbf{x}_r, \mathbf{x}_s, \omega) e^{-i\omega\left[\tau(\mathbf{x}_r, \mathbf{y}^S) + \tau(\mathbf{x}_s, \mathbf{y}^S)\right]}.$$

We then cross correlate these migrated traces

$$\begin{aligned} \mathcal{I}^{\mathrm{CINT}}(\mathbf{y}^S; \mathbf{w}, \widehat{f}) &= \int_{|\omega-\omega_0|\leq B} d\omega \int_{\substack{|\omega' - \omega_0| \leq B \\ |\omega' - \omega| \leq \Omega_d}} d\omega' \\ &\quad \sum_{r=1}^{N_r} \sum_{\substack{r' = 1 \\ |\mathbf{x}_r - \mathbf{x}'_r| \leq \frac{2c_0}{(\omega+\omega')\kappa_d}}}^{N_r} \widehat{m}(\mathbf{x}_r, \mathbf{y}^S, \omega) \overline{\widehat{m}(\mathbf{x}_{r'}, \mathbf{y}^S, \omega')}, \end{aligned} \qquad (1)$$

where $\mathbf{w} = (w_1, ..., w_{N_s})$.

The algorithm, continued.

We determine $\mathbf{w}$ and $\widehat{f}$ by minimizing an objective function $\mathcal{O}(\mathbf{w}, \widehat{f})$ that quantifies the quality of the image. We take it to be the L^1 norm. We then determine the weights $\mathbf{w} = (w_1, \ldots, w_{N_s})$ and the waveform $\widehat{f}(\omega)$ as minimizers of $\mathcal{O}(\mathbf{w}, \widehat{f})$, subject to the following constraints.

The weights should be nonnegative and sum to one

$$\sum_{s=1}^{N_s} w_s = 1, \quad w_s \geq 0, \quad s = 1, \ldots, N_s. \tag{2}$$

The support of $\widehat{f}(\omega) = \widehat{f}_B(\omega - \omega_0)$ is restricted to the fixed frequency band $[\omega_0 - B, \omega_0 + B]$, we ask that

$$\widehat{f}_B(\omega - \omega_0) \geq 0, \quad \text{for all } \omega \in [\omega_0 - B, \omega_0 + B] \tag{3}$$

and its integral is one.

Optimal subspace selection and CINT

Another way to introduce an optimization process using the SVD is by subspace selection. Let

$$D\left[\widehat{\Pi}(\omega);\omega\right] = \sum_{j=1}^{p} \sigma_j(\omega) d_j(\omega) \hat{\mathbf{u}}_j(\omega) \hat{\mathbf{v}}_j^*(\omega)$$

a subspace selector with weights $\{d_j(\omega)\}$. Now let

$$m(\mathbf{x}_r, \mathbf{x}_s, \mathbf{y}^S, \omega; d) = D\left[\widehat{\Pi}(\omega);\omega\right]_{rs} e^{-i\omega(\tau(\mathbf{x}_r,\mathbf{y}^S)+\tau(\mathbf{x}_s,\mathbf{y}^S))}$$

The CINT imaging functional to be optimized is:

$$I^{CINT}(\mathbf{y}^S; d) = \int\int_{|\omega-\omega'|\leq\Omega_d} d\omega d\omega' \sum_{|\mathbf{x}_r-\mathbf{x}_{r'}|\leq X_d} \sum_{|\mathbf{x}_s-\mathbf{x}_{s'}|\leq X_d}$$

$$m(\mathbf{x}_r, \mathbf{x}_s, \mathbf{y}^S, \omega; d) m(\mathbf{x}'_r, \mathbf{x}'_s, \mathbf{y}^S, \omega; d)$$

Summary of CINT and imaging in random media

- Coherent interferometry, which is the back propagation of local cross-correlations of traces, deals well with partial loss of coherence in cluttered environments.

- Adaptive estimation of the space-frequency decoherence addresses well the issue of learning the unknown environment.

- The key parameters Ω_d and X_d, which characterize the clutter, play a triple role: they are thresholding parameters for CINT, they determine its resolution, and characterize the coherence of the data. Theory and implementation issues merge.

- In Optimal Subspace and Illumination selection is computationally intensive but makes a huge difference.

Conclusions

Imaging in its many forms is at the center of modern applied mathematics.

- It is naturally interdisciplinary
- It is profoundly mathematical
- It has to deal with large data sets
- It has to deal with statistical issues
- It has to deal with optimization issues

Background References

1. 3D Seismic Imaging, B.L. Biondi, no 14 in Investigations in Geophysics (Tulsa: Society of Exploration Geophysics). Recent book on basic seismic imaging with detailed treatment of velocity analysis
2. Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion, Bleistein N, Cohen J K and Stockwell J W Jr, Springer, 2001. Mathematical treatment of migration imaging
3. Wave propagation and time reversal in randomly layered media, Fouque, Garnier, Papanicolaou and Solna, Springer, 2007. The first five chapters cover basic wave propagation in layered media. Chapter 6 is a self-contained treatment of asymptotics for stochastic equations as it is used in the rest of the book. Chapter 7 is a basic, mathematical treatment of waves in one-dimensional random media
4. Wave propagation and scattering in the heterogeneous Earth, Sato H and Fehler M 1998 (New York: Springer-Verlag). The use of radiative transport in seismology
5. Transport equations for elastic and other waves in random media, Ryzhik L V, Papanicolaou G C and Keller J B, 1996, Wave Motion, vol 24, 327-370. A more mathematical treatment of topics found in the previous reference
6. Imaging the Earth's interior, Claerbout J F, 1985, (Palo Alto: Blackwell Scientific Publications). Introduction to migration imaging

References specific to the Lectures

1. Theory and applications of time reversal and interferometric imaging, Borcea, Papanicolaou and Tsogka, Inverse Problems, vol 19, (2003), pp. 5139-5164. Contains details of many calculations in the slides (Lectures II,III, VII,VIII)
2. Statistical stability in time reversal, Papanicolaou, Ryzhik and Solna, SIAM J. on Appl. Math., 64 (2004), pp. 1133-1155. Contains details on calculations in Lectures VII-VIII
3. Self-averaging from lateral diversity in the Ito-Schroedinger equation, George Papanicolaou, Leonid Ryzhik and Knut Solna, SIAM Journal on Multiscale Modeling and Simulation, vol 6, (2007), pp. 468-492. Contains details (more mathematical) of lectures VII-VIII
4. Imaging and time reversal in random media, Liliana Borcea, Chrysoula Tsogka, G. Papanicolaou and James Berryman, Inverse Problems, 18 (2002), pp. 1247–1279. Use of SVD in imaging as in Lecture V
5. Edge illumination and imaging of extended reflectors, Liliana Borcea, George Papanicolaou and Fernando Guevara Vasquez, SIAM Journal on Imaging Sciences, vol 1 (2008), pp. 75-114. Use of SVD in edge illumination and imaging as in Lecture VI
6. Passive Sensor Imaging Using Cross Correlations of Noisy Signals in a Scattering Medium, Josselin Garnier and George Papanicolaou in: http://math.stanford.edu/~papanico. The paper on which Lecture IV is based, as well as the Lecture of Garnier

References specific to the Lectures, continued

7 Stable iterative reconstruction algorithm for nonlinear travel time tomography Berryman J, 1990, Inverse Problems, vol 6, 21-42. Basic reference for travel time tomography

8 Adaptive interferometric imaging in clutter and optimal illumination Borcea, Papanicolaou, and Tsogka, 2006, Inverse Problems, vol 22, 1405-1436. Paper on which Lectures IX and X are based, plus Tsogka's lectures

9 Optimal illumination and waveform design for imaging in random media Borcea, Papanicolaou, and Tsogka, 2007, J. Acoust. Soc. Am., vol 122, 3507-3518. Same as previous reference

References, on Green's function from cross correlations

1. Surface wave tomography from microseisms in Southern California, Sabra K G, Gerstoft P, Roux P, and Kuperman W, 2005, Geophys. Res. Lett., vol 32, L14311
2. Interferometric daylight seismic imaging, Schuster G T, Yu J, Sheng J and Rickett J, 2004, Geophysical Journal International, vol 157, 832-852
3. Velocity inversion by differential semblance optimization, Symes W W and Carazzone J J, 1991, Geophysics, vol 56, 654-663
4. Green's function representations for seismic interferometry, Wapenaar K and Fokkema J, 2006, Geophysics, vol 71, SI33-SI46
5. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis I. Phase velocity maps, Yao H, van der Hilst R D, and de Hoop M V, 2006, Geophysical Journal International, vol 166, 732-744
6. Ambient noise cross correlation in free space: Theoretical approach, Roux P, Sabra K G, Kuperman W A, and Roux, A 2005, J. Acoust. Soc. Am., 117, 79-84

References for cross correlations, continued

7 Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring, Rickett J and Claerbout J, 1999, The Leading Edge, vol 18, 957-960

8 Correlation of random wave fields: an interdisciplinary review, Larose E, Margerin L, Derode A, Van Tiggelen B, Campillo M, Shapiro N, Paul A, Stehly L, and Tanter M, 2006, Geophysics, vol 71, SI11-SI21

9 Seismic interferometry - turning noise into signal, Curtis A, Gerstoft P, Sato H, Snieder R, and Wapenaar K, 2006, The Leading Edge, vol 25, 1082-1092