Mathematical Problems in Geophysical Wave Propagation

George Papanicolaou

Abstract. We review several aspects of the mathematical theory of wave propagation in random media with particular emphasis on topics of geophysical interest.

Contents

1 Introduction

1.1 Why geophysical wave propagation

In deciding what to present as a plenary lecture in applied mathematics and probability at ICM-98, I considered several areas with which I am familiar and decided to focus on geophysical wave propagation for a couple of reasons. One, a technical one, is that inhomogeneities are strong and highly anisotropic so that the modeling and analysis of wave propagation in the earth's crust is mathematically interesting and quite difficult. The other reason is closely related to what my view of modern applied mathematics is: the creation and development of a mathematical environment for physical, economic, biological or other phenomena. This involves active participation of the mathematician in the quantitative modeling, in the analysis, in the computations, as well as in the interpretation of results and assessment of the effectiveness of the modeling.

The resulting mathematical methodology will be uneven, from routine off-the-shelf toolbox applications to entirely uncharted problems that need new ideas and techniques, and it is up to the mathematician to decide what the right mix of mathematical sophistication and rough heuristics should be. An overly mathematical approach will impede communication with nonmathematical specialists who value results and do not care much for mathematical generality. Accepting the conventional wisdom in a field, and concentrating on technical mathematical issues, is not a good idea either. Geophysical wave propagation is a case in point. It is fair to say that wave localization is virtually unknown to geophysicists. But, as I will try to explain in this lecture, wave localization is quite important in exploration geophysics because, among other things, it influences the resolution of seismic imaging and the effective depth penetration of seismic probes. What is the best way to approach these problems mathematically?

A few years ago, K. Aki, a distinguished seismologist whose ideas about the role of crustal inhomogeneities in seismic wave propagation have been very influential, heard a seminar that I gave on wave localization and asked this question: How can one tell from seismic observations that wave localization has taken place? Electronic wave localization in semiconductors goes back forty years [1], with the strong participation of mathematicians during the last twenty years, so we should be able to say quite a bit, as I will try to explain in this lecture. But Aki's question is a profound one that leads to the most complex and least understood issue in geophysical wave propagation, the localization-transport transition. It is a pragmatic, operational question which reminds us that great intellectual challenges can have humble, unpretentious origins. I think that it takes a mathematician to answer Aki's question and perhaps it will be one that does it.

1.2 Random media or environments in general

I will treat the earth's crust as a random medium, that is, as an elastic medium with density and Lame parameters that are random functions of space. The equations of linear elastic wave propagation become now stochastic partial differential equations. Initial and boundary conditions must also be specified and they could bring in additional randomness, from modeling the rough surface of the earth. At this level of generality the randomness is nothing more than variable coefficients and non-flat boundaries, so general linear PDE methods can deal with everything (symmetric hyperbolic systems). If dissipation is important, and it is in some contexts, it can be put into the equations in different ways. There is no general agreement on how to best model dissipation analytically and this is an interesting issue that I will not address here.

But even this much is somewhat grudgingly accepted by geophysicists. I regularly hear comments like: there is only one earth and it is not changing all that fast, so where is the statistical ensemble of realizations coming from? If stochastic modeling is to be criticized along such lines then why are we modeling the Dow Jones Industrial Average, or some other index or asset price, as a stochastic process? There is only one realization of the DJIA just as there is only one realization of the earth. What the stochastic processes model is uncertainty, lack of information and its consequences when only imperfect and sparse observations or measurements are available, and even desirable. The notion of 'effective' medium is very much part of the mathematical physics of the 19th century, of Maxwell, Rayleigh and others, which is why equations with constant coefficients have any relevance at all in modeling. The conceptual barrier seems to come up when one thinks of fluctuations.

It is not an accident, therefore, that in one of the first instances of wave propagation in random media, natural light propagation through a turbulent atmosphere, astrophysicists at the turn of the 20th century did not go to Maxwell's equations (or the wave equation if the vector nature of light waves can be neglected) but developed a new, phenomenological theory, the radiative transport theory, to interpret observed phenomena. There are a few isolated attempts to consider random media, with fluctuations, during the first half of the 20th century but it is with the advent of radar and sonar during in the forties that random waves emerge as a subject. Keller's papers in the sixties [2] where very influential because they were the first ones written by a mathematician, who thought about the conceptual foundations and separated heuristics from legitimate calculations. It was also in the sixties that the connection between radiative transport theory and stochastic wave equations was clarified, as I will discuss in section 3.1.

Atmospheric wave propagation, from radio to radar to optical frequencies, and underwater sound propagation, from 20 hertz to kilohertz, were the main applications driving the theory of wave propagation in random media in the seventies and are discussed in Ishimaru's book [3]. It is interesting to note that the notion of wave localization is nowhere to be found in this book. Random media in seismology appeared first in the mid eighties in a simple version of radiative transport [4]. Transport theory is now just beginning to become mainstream in seismology as is seen from the recent book of Fehler and Sato [5]. But wave localization is not discussed in this book either. A treatment of waves in random media that deals extensively with wave localization is given by Ping Sheng [6].

What is wave localization anyway? I will explain it in some detail in section 4.3 but, roughly, it is when random inhomogeneities trap wave energy in a finite region and do not allow it to spread as it would normally. Random media behave then like periodic media that have band-gap spectra, allowing wave propagation in some frequency ranges but not in others. It is remarkable that this happens for random media that are not close to periodic ones at all. Mathematically, it is shown that wave or wave-like operators with stationary (translation invariant) random coefficients in unbounded regions have discrete spectra [7]. Discrete spectra means that the wave energy in each mode initially will remain there for ever, oscillating in time but not propagating out to infinity. In three dimensional wave propagation this can happen only when parameter fluctuations are very large. This is not the case for electromagnetic waves in the atmosphere or sound waves in the ocean. The fluctuations are weak, a few percent, and when they are important they lead to radiative transport, which allows spreading of wave energy in diffusive rather than wave-like manner.

Where then do we see wave localization in classical wave propagation? We see it when wave energy is channeled, by a waveguide, by a transmission line, by an optical fiber, by strong anisotropy due to layering in the lithosphere, etc. We also see it in nearly periodic structures. Waves in an one dimensional random medium will localize, even if the fluctuations in the medium parameters are weak. In geophysical wave propagation and elsewhere (in optical localization) a key issue is the identification of structures, more complicated than simple channeling or periodicity, that tend to enhance the onset of wave localization by random fluctuations. This is the localization-transport transition problem.

1.3 Acknowledgements

I would like to thank several friends and colleagues with whom I worked in the past: R. Burridge, J.B. Keller, W. Kohler and B.S. White with whom I have been associated for a long time, and my more recent collaborators: M. Asch, G. Bal, J. Berryman, J. Bronski, O. Dorn, J.P. Fouque, F. Herrmann, R. Knapp, P. Lewicki, A. Nachbin, M. Postel, L. Ryzhik, Y. Samuelides, P. Sheng, K. Solna and S. Weinryb.

2 General notions about waves in random media

2.1 Scales

There are three basic length scales in wave propagation phenomena:

- The typical wavelength λ
- The typical propagation distance L
- The typical size of the inhomogeneities l

In geophysical wave propagation it is difficult to associate a 'typical' scale that is characteristic of the inhomogeneities. The density and local speed of propagation of waves vary on many scales. We may think of l as a typical correlation length. When the standard deviation of the fluctuations is small then the most effective interaction of the waves with the random medium will occur when $l \sim \lambda$, that is, the wavelength is comparable to the correlation length. And this interaction will not be observable unless the propagation distance is large ($L >> \lambda$). If propagation distances are short, a few wavelengths or correlation lengths, then effective medium theories will work fine. There will be a deterministic propagation speed (for scalar waves), the effective speed, with which energy will propagate as if the medium were deterministic. The effective medium theory will be valid also when the wavelength is long compared to propagation distances, even if the correlation length is short ($l << L$) and the fluctuations have a large standard deviation. This is the homogenization limit.

Of course this rough way of thinking with scales does not capture the effect of a waveguide geometry, or the effective dimension of the propagation phenomenon. But thinking with scales is very useful and, with some experience, it can become a very good heuristic tool.

2.2 Types of waves

It is classical waves, solutions of the wave equation or more general symmetric hyperbolic systems, that we want consider, rather than electronic waves which are solutions of the Schrödinger equation. The waves are vector fields in general, as with electromagnetic waves which are solutions of Maxwell's equations or elastic waves where the elastic displacement field is a solution of the elastic wave equations. Mode conversion, the transfer of energy from compression to shear waves for example, is an important effect in random media. So is polarization, which is associated with vector waves all of whose components travel with the same speed. Polarization tends to get lost in a random medium and the way this happens is an important way to make inferences about the nature of the propagation environment.

2.3 Coherent and incoherent fields

When the random fluctuations of the medium parameters are small then the random fluctuations in the solutions will be small, if the propagation distances are not too big. The mean solution, the coherent field, will carry most of the energy. As the waves propagate their fluctuating component, the incoherent field, gets more energy. The total energy is conserved, if there is no dissipation, but the coherent field loses energy and slows down. This behavior of the coherent field is something that can be calculated easily and is well established in the engineering literature.

2.4 Localization and transport

If fluctuations are weak and propagation distances large, most of the wave energy will be incoherent. In seismology, for example, after the first arrival from a disturbance far away the seismogram is dominated by strong fluctuations from multiple scattering. The later part, the coda of a seismogram is mostly incoherent field measurement. It is in this regime that radiative transport is a good approximation. It allows accurate calculation of the envelopes of the seismograms without resolving the detailed fluctuations. A new scale enters the description of propagation phenomena: the mean free path. This is a length scale that gives an indication of the importance of multiple scattering and is much more relevant than the correlation length of the inhomogeneities.

Wave localization is total trapping of the wave energy by scattering from the random inhomogeneities. It is the regime where fluctuation phenomena dominate so we have little intuition for what should happen. For one thing the random fluctuations must be very strong and the structure of the propagating medium must be special (a channeling medium or an ordered, periodic structure). In the lithosphere fluctuations in the speeds of propagation of elastic waves can be as large as 15% and they can be highly anisotropic, with horizontal correlation lengths much larger than vertical ones. Localization manifests itself in fat codas of seismograms, or codas with envelopes that decay slowly. This is a clear indication that there is a lot of multiple scattering going on. Moreover, radiative transport would tend to underestimate the size of the codas indicating that a different analytical theory is needed. What is missing at present is a robust and effective criterion for discriminating between these two situations.

2.5 Nonlinearity and randomness

Nonlinearity and randomness interact significantly only in very special situations, as in soliton propagation in optical fibers or when high intensity laser beams interact with material inhomogeneities. Nonlinearity is rarely an issue in seismic wave propagation except very near sources. In one dimensional wave propagation both nonlinearity and randomness are strongly felt and a long-standing problem is the analysis of their interaction. Is there, for example, wave localization when we have nonlinearities? This is a very difficult question that cannot be answered by a yes or no. The phenomena depend sensitively on the exact setup of the problem: the form of the nonlinearity, the various scales associated with the inhomogeneities and the propagation phenomenon, and the form of the excitation [8, 9, 10, 11, 12, 13].

2.6 Numerical simulations

At the dawn of the 21st century, when computational power is doubling every two years or so, and computational cost is dropping to the point where a good laptop computer today is more powerful than the Cray I supercomputer of the late seventies, why is anybody interested in analytical methods? We have the computational power to simulate anything we want and we have the ability to make detailed and extensive measurements, which in seismology result in huge data sets. What could mathematical analysis contribute in this context?

Being skeptical about the utility of mathematical analysis and believing that we can compute or simulate everything we need may appear naive to a mathematician but it is increasingly the dominant view in many fields, in geophysical wave propagation for example.

The fact is that if we want to understand the behavior of seismic codas we cannot rely on direct numerical simulations. If the typical wavelength is of the order of 3-5 km and we want to calculate a synthetic seismogram 1000 km from the source we need a spatial grid that has at least five points per wavelength, and more if we want to simulate accurately random fluctuations in the parameters. In a realistic three dimensional setup it is impossible to generate numerical solutions that will yield a 3 second synthetic seismogram with millisecond resolution. What is even more important to realize is that we should not really want to do this because with radiative transport theory seismic coda envelopes can be calculated. What is holding up realistic numerical computations is not computing power but analysis: we do not have good enough transport theoretic boundary conditions on the earth's surface and at interfaces. The mean free path may be as large as 20-30 km and Monte Carlo methods can give reasonably accurate solutions using a high-end workstation. Transport theory does what is called 'sub-grid' modeling in computational fluid dynamics. We do not have to resolve the small scale inhomogeneities if we can do some analysis, which is in fact difficult but doable.

2.7 Parameter estimation and imaging

Imaging of the earth's interior is a challenge that will be with us for a very long time because the inhomogeneities are so strong. In exploration seismology, where seismic probing can generate huge data sets, the issue is not so much good algorithms for imaging but low complexity algorithms. Efficient compression of geophysical data sets is perhaps the most urgent problem that exploration seismology faces at present.

It appears at first that this has nothing to do with waves. Wavelets or other tools for compression from signal processing come to mind, and they are being used. If noise effects are ignored and if the typical wave length of a probing pulse is 100-150 m (for shorter wavelengths dissipation effects are much stronger), we cannot expect image resolution better that 25 m or so at a depth of a few kilometers. And if noise and multiple scattering are to be taken into consideration it is not at all clear what the achievable resolutions are without some compensation. Noise reducing methods (stacking) that are used in imaging are not so effective. Much more needs to be done analytically here. Imaging itself, without noise, is based on variants of a backward wave propagation method (migration) that has now a substantial theoretical basis [14, 15, 16].

The best compression method is to go from the seismic data to the image itself, of course, so good compression has to be adapted to the specific data set and its structure. But there must be interesting algorithms, yet to be found, that are somewhere between know-nothing methods like wavelet decomposition and thresholding, and know-all full imaging.

3 The transport regime

Radiative transport is a phenomenological theory that was introduced to describe the propagation of light intensity through the Earth's atmosphere. It has been applied successfully to many other problems of wave propagation in a complex medium. In its simplest form, let $a(t, \mathbf{x}, \mathbf{k})$ denote the angularly resolved energy density defined for all wave vectors $\mathbf{k}$, position $\mathbf{x}$ and time t. Because of interaction with the inhomogeneous medium through which it propagates, a wave with wave vector $\mathbf{k}$ may be scattered into any other direction $\hat{\mathbf{k}}'$, where $\hat{\mathbf{k}} = \frac{\mathbf{k}}{|\mathbf{k}|}$. The transport equation gives the energy balance

$$\frac{\partial a(t,\mathbf{x},\mathbf{k})}{\partial t} + \nabla_{\mathbf{k}}\omega(\mathbf{x},\mathbf{k}) \cdot \nabla_{\mathbf{x}} a(t,\mathbf{x},\mathbf{k}) - \nabla_{\mathbf{x}}\omega(\mathbf{x},\mathbf{k}) \cdot \nabla_{\mathbf{k}} a(t,\mathbf{x},\mathbf{k}) = \int_{R^n} \sigma(\mathbf{x},\mathbf{k},\mathbf{k}')a(t,\mathbf{x},\mathbf{k}')d\mathbf{k}' - \Sigma(\mathbf{x},\mathbf{k})a(t,\mathbf{x},\mathbf{k}). \tag{1}$$

Here n is the dimension of space ($n = 2$ or 3), $\omega(\mathbf{x},\mathbf{k})$ is the local frequency at position $\mathbf{x}$ of the wave with wave vector $\mathbf{k}$, the differential scattering cross-section $\sigma(\mathbf{x},\mathbf{k},\mathbf{k}')$ is the rate at which energy with wave vector $\mathbf{k}'$ is converted to wave energy with wave vector $\mathbf{k}$ at position $\mathbf{x}$, and

$$\int \sigma(\mathbf{x},\mathbf{k}',\mathbf{k})d\mathbf{k}' = \Sigma(\mathbf{x},\mathbf{k}) \tag{2}$$

is the total scattering cross-section. The function $\sigma(\mathbf{x},\mathbf{k},\mathbf{k}')$ is nonnegative and usually symmetric in $\mathbf{k}$ and $\mathbf{k}'$. The left side of (1) is the total time derivative of $a(t,\mathbf{x},\mathbf{k})$ at a point moving along a trajectory in phase space $(\mathbf{x},\mathbf{k})$ and may be written as a Liouville equation

$$\frac{\partial a}{\partial t} = \{\omega, a\}, \tag{3}$$

where $\{f,g\} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\frac{\partial g}{\partial k_i} - \frac{\partial f}{\partial k_i}\frac{\partial g}{\partial x_i} \right)$ is the Poisson bracket. The right side of (1) represents the effects of scattering.

The transport equation (1) is conservative when (2) holds because then

$$\iint a(t,\mathbf{x},\mathbf{k})d\mathbf{x}d\mathbf{k} = \mathrm{const}$$

independent of time. Absorption may be accounted for easily by letting the total scattering cross-section be the sum of two terms

$$\Sigma(\mathbf{x},\mathbf{k}) = \Sigma_{sc}(\mathbf{x},\mathbf{k}) + \Sigma_{ab}(\mathbf{x},\mathbf{k})$$

where $\Sigma_{sc}(\mathbf{x},\mathbf{k})$ is the total scattering cross-section given by (2) and $\Sigma_{ab}(\mathbf{x},\mathbf{k})$ is the absorption rate.

The radiative transport equation (1) was derived from the microscopic equations in the sixties and seventies by many authors (see [17] for references). A nice overview of these methods and results is presented in a recent review [18]. We have recently considered scattering of high frequency waves in a random medium [17] and established validity of the radiative transport theory for scalar and vector waves, including mode conversion and polarization in the following regime:

- Distances of propagation L are much larger than the wave length λ,

- The medium parameters vary on the scale comparable to the wave length,
- The mismatch between the inhomogeneities and the background medium is small,
- Absorption is small.

This regime arises in many physically important situations. In seismic wave propagation, teleseismic events can be modeled by radiative transport equations [4, 5].

3.1 WAVES TO TRANSPORT

Transport equations for the phase space wave energy densities are constructed [17, 19, 20] as follows. We assume here that the space domain is R^3 ($n = 3$) and deal with acoustic waves. The acoustic equations for the velocity $\mathbf{v}$ and pressure p are

$$\rho \frac{\partial \mathbf{v}}{\partial t} + \nabla p = 0 \tag{4}$$
$$\kappa \frac{\partial p}{\partial t} + \nabla \cdot \mathbf{v} = 0.$$

This system may be written in a general form of a symmetric hyperbolic system (with convention of summation over repeated indices):

$$A(\mathbf{x}) \frac{\partial \mathbf{u}}{\partial t} + D^j \frac{\partial \mathbf{u}}{\partial x^j} = 0, \tag{5}$$

where $\mathbf{u} = (\mathbf{v}, p)$, and $\mathbf{x} \in R^n$. The matrix $A(\mathbf{x}) = \mathrm{diag}(\rho, \rho, \rho, \kappa)$ is symmetric and positive definite and the matrices D^j are symmetric and independent of $\mathbf{x}$ and t. We consider high frequency solutions of (5). Physically this means that the typical wave length λ of the initial data is much smaller than the overall propagation distance L with $\varepsilon = \dfrac{\lambda}{L} \ll 1$. The spatial energy density for the solutions of (5) is given by

$$\mathcal{E}(t, \mathbf{x}) = \frac{\rho \mathbf{v}^2}{2} + \frac{\kappa p^2}{2} = \frac{1}{2}(A(\mathbf{x})\mathbf{u}(t, \mathbf{x}) \cdot \mathbf{u}(t, \mathbf{x})) = \frac{1}{2} A_{ij}(\mathbf{x}) u_i(t, \mathbf{x}) \bar{u}_j(t, \mathbf{x}) \tag{6}$$

and the flux $\boldsymbol{\mathcal{F}}(\mathbf{x})$ by

$$\boldsymbol{\mathcal{F}}_i(t, \mathbf{x}) = p\mathbf{v} = \frac{1}{2}(D^i \mathbf{u}(t, \mathbf{x}) \cdot \mathbf{u}(t, \mathbf{x})). \tag{7}$$

We have the energy conservation law

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot \boldsymbol{\mathcal{F}} = 0. \tag{8}$$

and thus the total energy is conserved:

$$\frac{d}{dt} \int \mathcal{E}(t, \mathbf{x}) d\mathbf{x} = 0. \tag{9}$$

The high frequency limit $\varepsilon \to 0$ of the energy density $\mathcal{E}(t,\mathbf{x})$ is described in terms of the Wigner transform, which is defined by

$$W_\varepsilon(t,\mathbf{x},\mathbf{k}) = \left(\frac{1}{2\pi}\right)^n \int e^{i\mathbf{k}\cdot\mathbf{y}} \mathbf{u}_\varepsilon(t,\mathbf{x}-\varepsilon\mathbf{y}/2)\mathbf{u}_\varepsilon^*(\mathbf{x}+\varepsilon\mathbf{y}/2)d\mathbf{y}, \tag{10}$$

where $\mathbf{u}_\varepsilon(t,\mathbf{x})$ is the solution of (5). The Wigner transform W_ε is a 4×4 Hermitian matrix. Its limit as $\varepsilon \to 0$ is called the Wigner distribution and is denoted by $W(t,\mathbf{x},\mathbf{k})$. The limit Wigner matrix is not only Hermitian but also positive definite. The limit energy density and flux are expressed in terms of $W(t,\mathbf{x},\mathbf{k})$ by

$$\mathcal{E}(t,\mathbf{x}) = \frac{1}{2}\int \mathrm{Tr}(A(\mathbf{x})W(t,\mathbf{x},\mathbf{k}))d\mathbf{k}$$

and

$$\mathcal{F}_i(t,\mathbf{x},\mathbf{k}) = \frac{1}{2}\int \mathrm{Tr}(D^i W(t,\mathbf{x},\mathbf{k}))d\mathbf{k}.$$

The limit Wigner distribution may be decomposed over different wave modes in a way that generalizes the plane wave decomposition in a homogeneous medium. The dispersion matrix of the system (5) is defined by

$$L(\mathbf{x},\mathbf{k}) = A^{-1}(\mathbf{x})k_i D^i = \begin{pmatrix} 0 & 0 & 0 & k_1/\rho \\ 0 & 0 & 0 & k_2/\rho \\ 0 & 0 & 0 & k_3/\rho \\ k_1/\kappa & k_2/\kappa & k_3/\kappa & 0 \end{pmatrix}. \tag{11}$$

It has one double eigenvalue $\omega_1 = \omega_2 = 0$ and two simple eigenvalues

$$\omega_f = v|\mathbf{k}| \ , \quad \omega_b = -v|\mathbf{k}| \ , \tag{12}$$

where $|\mathbf{k}| = \sqrt{k_1^2+k_2^2+k_3^2}$ and v is the sound speed

$$v = \frac{1}{\sqrt{\kappa\rho}}. \tag{13}$$

The corresponding basis of eigenvectors is

$$\begin{aligned} \mathbf{b}_1 &= \frac{1}{\sqrt{\rho}}(\mathbf{z}^{(1)}(\mathbf{k}),0)^t, \quad \mathbf{b}_2 = \frac{1}{\sqrt{\rho}}(\mathbf{z}^{(2)}(\mathbf{k}),0)^t, \\ \mathbf{b}_f &= (\frac{\hat{\mathbf{k}}}{\sqrt{2\rho}}, \frac{1}{\sqrt{2\kappa}})^t, \quad \mathbf{b}_b = (\frac{\hat{\mathbf{k}}}{\sqrt{2\rho}}, -\frac{1}{\sqrt{2\kappa}})^t, \end{aligned} \tag{14}$$

where the vectors $\hat{\mathbf{k}}$, $\mathbf{z}^{(1)}(\mathbf{k})$ and $\mathbf{z}^{(2)}(\mathbf{k})$ form an orthonormal triplet. The eigenvectors $\mathbf{b}_1(\mathbf{k})$ and $\mathbf{b}_2(\mathbf{k})$ correspond to transverse advection modes, orthogonal to the direction of propagation. These modes do not propagate because $\omega_{1,2} = 0$. The eigenvectors $\mathbf{b}_f(\mathbf{k})$ and $\mathbf{b}_b(\mathbf{k})$ represent forward and backward acoustic waves, which are longitudinal , and which propagate with the sound speed v given by (13).

The limit Wigner distribution matrix $W(t, \mathbf{x}, \mathbf{k})$ has the form [17]:

$$\begin{aligned} W(t, \mathbf{x}, \mathbf{k}) =& \sum_{\tau=1}^{2} W_{ij}^{\tau}(t, \mathbf{x}, \mathbf{k}) \mathbf{b}^{i}(\mathbf{k}) \mathbf{b}^{j*}(\mathbf{k}) \\ &+ a_f(t, \mathbf{x}, \mathbf{k}) \mathbf{b}_f(\mathbf{k}) \mathbf{b}_f^{*}(\mathbf{k}) + a_b(t, \mathbf{x}, \mathbf{k}) \mathbf{b}_b(\mathbf{k}) \mathbf{b}_b^{*}(\mathbf{k}). \end{aligned} \tag{15}$$

The first term corresponds to the non-propagating modes and may be set to zero here without any loss of generality. The last two terms correspond to forward and backward propagating sound waves. The scalar functions $a_{f,b}$ are related by $a_f(t, \mathbf{x}, \mathbf{k}) = a_b(t, \mathbf{x}, -\mathbf{k})$, and a_f satisfies the Liouville equation

$$\frac{\partial a}{\partial t} + \nabla_{\mathbf{k}} \omega \cdot \nabla_{\mathbf{x}} a - \nabla_{\mathbf{x}} \omega \cdot \nabla_{\mathbf{k}} a = 0. \tag{16}$$

They may be interpreted as phase space energy densities since they are non-negative (because the matrix $W(t, \mathbf{x}, \mathbf{k})$ is non-negative) and

$$\mathcal{E}(\mathbf{x}) = \frac{1}{2} \int d\mathbf{k} [a_f(t, \mathbf{x}, \mathbf{k}) + a_b(t, \mathbf{x}, \mathbf{k})] = \int d\mathbf{k} a_f(t, \mathbf{x}, \mathbf{k}).$$

The flux is given by

$$\mathcal{F} = \frac{v}{2} \int d\mathbf{k} [\hat{\mathbf{k}} a_f(t, \mathbf{x}, \mathbf{k}) - \hat{\mathbf{k}} a_b(t, \mathbf{x}, \mathbf{k})] = v \int d\mathbf{k} \hat{\mathbf{k}} a_f(t, \mathbf{x}, \mathbf{k}). \tag{17}$$

The radiative transport equation (1) arises when the density ρ and compressibility κ are random and oscillating on the scale of the wave length, so we assume they have the form

$$\rho \to \rho(1 + \sqrt{\varepsilon} \rho_1(\frac{\mathbf{x}}{\varepsilon})), \quad \kappa \to \kappa(1 + \sqrt{\varepsilon} \kappa_1(\frac{\mathbf{x}}{\varepsilon})).$$

The random processes ρ_1 and κ_1 are mean zero space homogeneous with power spectral densities $\hat{R}_{\rho\rho}$, $\hat{R}_{\kappa\kappa}$, and cross spectral density $\hat{R}_{\kappa\rho}$. The limit $\varepsilon \to 0$ is the high frequency limit since the parameter ε is the ratio of wave length to propagation distance. In (3.1) we take the ratio of correlation length to propagation distance to be of order ε also, and we take the standard deviation of the fluctuations to be of order $\sqrt{\varepsilon}$. It is in this scaled limit that radiative transport theory emerges. The radiative transport equation for $a(t, \mathbf{x}, \mathbf{k}) = a_f(t, \mathbf{x}, \mathbf{k})$ is

$$\begin{aligned} \frac{\partial a}{\partial t} + v \hat{\mathbf{k}} \cdot \nabla_{\mathbf{x}} a - |\mathbf{k}| \nabla_{\mathbf{x}} v \cdot \nabla_{\mathbf{k}} a =& \frac{\pi v^2 |\mathbf{k}|^2}{2} \int \delta(v|\mathbf{k}| - v|\mathbf{k}'|)[a(\mathbf{k}') - a(\mathbf{k})] \\ &\cdot \left\{ (\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}')^2 \hat{R}_{\rho\rho}(\mathbf{k} - \mathbf{k}') + 2(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}') \hat{R}_{\rho\kappa}(\mathbf{k} - \mathbf{k}') + \hat{R}_{\kappa\kappa}(\mathbf{k} - \mathbf{k}') \right\} d\mathbf{k}'. \end{aligned} \tag{18}$$

This equation is of the form (1). The *mean free path* is a typical value of the ratio $\frac{v}{\Sigma}$, the speed over the total scattering cross-section. It can be thought of as the distance over which scattering by the inhomogeneities is effective. It is a length scale that can be estimated from seismic data while correlation lengths and standard deviations of parameter fluctuations are usually not observable.

The radiative transport equation (1) has been derived from equations governing particular wave motions by various authors, such as Stott [21], Watson et.al. [22], [23], [24], Barabanenkov et.al. [25], Besieris and Tappert [26], Howe [27], Ishimaru [3] and Besieris et. al. [28] with a recent survey presented in [29]. These derivations also determine the functions $\omega(\mathbf{x},\mathbf{k})$ and $\sigma(\mathbf{x},\mathbf{k},\mathbf{k}')$ and show how a is related to the wave field. In [17], (1) and these functions are derived as a special case of a more general theory using the Wigner distribution and symmetric hyperbolic systems.

There is not a lot of mathematical work on the wave-to-transport limit, and most of it is for the Schrödinger equation with random potential. We cite here the work of Martin and Emch [30], of Spohn [31], of Dell'Antonio [32] and the recent extensive study of Ho, Landau and Wilkins [33] as well as [34]. They treat only spatially homogeneous problems but it is known how to extend the analysis to the spatially inhomogeneous case (slow x-dependent initial data and potential) [35]. A really satisfactory mathematical treatment of radiative transport asymptotics from random wave equations is lacking at present.

3.2 Transport for electromagnetic and elastic waves

Transport theory for electromagnetic and elastic waves is interesting because of wave polarization. This is important in astrophysics and is analyzed in great detail in Chandrasekhar's treatise [36]. Coherence of polarized light persists and must be tracked, leading to a *system* of transport equations for the Stokes parameters that fix the state of polarization. The derivation of this system from Maxwell's equations was first done in the early seventies, and using symmetric hyperbolic systems and Wigner distributions in [17], where the earlier papers are cited.

The main reason we wanted a general derivation of transport equations for general waves was so that we could deal with elastic waves. One can, of course, write down phenomenological equations for the transport of elastic wave energy and this was done often in the last 10-15 years [5]. The problem is that shear waves were treated like acoustic waves and the role of polarization was not accounted for correctly in the geophysics literature, even though the similarity with electromagnetic waves (Chandrasekhar's work) should be clear. In [17] it is shown that elastic wave transport is like E&M for shear and like acoustics for compressional waves, and the two wave modes are coupled.

A simple but interesting consequence of the general derivation is the symmetry (self-adjointness) of the transport equations. This implies immediately that the only equilibrium phase space energy densities are the uniform ones (over the support of the energy surface). The spatial energy densities for the compressional P waves and the shear S waves must be in a fixed ratio to each other, which turns out to be

$$\mathcal{E}_P = \frac{v_S^3}{2v_P^3}\mathcal{E}_S$$

Here v_S is the shear speed (about 3km/sec) and v_P is the compressional speed (about 5km/sec). This makes the P wave energy about one tenth of the S wave energy deep in the coda of seismograms, assuming surface effects are not important so that the free space theory can be used. This is independent of what the source is and of the details of the scattering medium, as long as there is effective scattering. The asymptotic energy law is well known empirically but it did not have an explanation from first principles so we presented it in detail and related it to the seismological literature in [37]. It turns out that this kind of long time P-to-S energy equilibrium was known in connection with remote sensing with ultrasound [41].

3.3 Boundary conditions

Finding appropriate transport theoretic boundary conditions for wave propagation in the transport regime is perhaps the most pressing issue both theoretically and for the applications, in geophysics, in electromagnetics, in ultrasound and elsewhere. The problems are analytically difficult as can be seen from [20] where the relatively simple case of inhomogeneous, slowly varying deterministic media with a flat interface is considered and transport theoretic boundary conditions are derived in the high frequency limit.

There is a lot of physical and applied literature on scattering form random rough surfaces [38, 39]. The issue is to determine what is appropriate as a boundary or interface condition for radiative transport equations. As with polarization, interfaces are a source of coherence in an otherwise incoherent scattering process. So they must be treated carefully to avoid oversimplifications. In [40] we consider acoustic reflection and transmission by a flat interface and derive transport theoretic boundary conditions, but a lot more has to be done here, including the derivation of boundary conditions for E&M and elastic wave transport.

3.4 The diffusive regime

It is well known, primarily from studies that originated in neutron scattering and reactor theory, that when the propagation distance in the transport regime is large compared to the mean free path a simpler diffusion theory emerges. In some seismic propagation problems the mean free path is 20-30 km but propagation is over 1000 km and more. So it is quite clear that a diffusion approximation for the transport equations is called for. We know how to do this when there are no boundaries present [17], even with polarization for E&M and elastic waves.

The problem is that the crustal wave guide is 30-40 km deep and it is not clear how to use the diffusion approximation, or even how to decide if it should be used at all. But the mathematical problem of finding asymptotic boundary and interface conditions in the diffusive regime is interesting, quite delicate analytically and potentially very useful [42]. In radar scattering from clutter, the diffusive transport theory is very likely the most appropriate one to use for wavelengths in the 10 cm to 1 m range, for example.

3.5 Parameter identification and inverse problems

Parameter identification for radiative transport has received relatively little attention in geophysics [5]. In light propagation through the atmosphere the situation is, of course, very different if only because the measurements that can be made are very different. The recent activity in diffusive tomography [43] should eventually find applications in geophysics as well, but there are many difficult problems that must be settled along the way, such as getting the right transport theoretic boundary conditions.

4 The localization regime

I will review briefly reflection of acoustic plane wave pulses normally incident on a randomly layered half space, $z < 0$, with z the direction of the layering [45, 46]. A good reference for deterministic wave propagation in layered media is Brekhovskikh's book [44]. It is in randomly layered media that wave localization is dominant. I will describe it in the time domain, for pulses, because this is the most interesting case in geophysical wave propagation, in reflection seismology and elsewhere. It is also not treated much in the mathematical or physical literature specialized to localization problems, and the simple intuition that most specialists have for time harmonic, one dimensional wave localization is not quite adequate for pulses. This was pointed out some time ago [50].

Radiative transport theory is not, of course, valid for randomly layered media. This was also considered long ago in connection with wave guides and optical fibers [47]. But it is not well understood in applied fields, even today, and papers appear occasionally that attempt to 'derive' radiative transport equations for propagation in layered media. I do not mean here three dimensional radiative transport in plane parallel structures. I mean random layering. If radiative transport were valid in this case, then the differential scattering cross-section would be singular, concentrated in only two (in the simplest case) directions, up and down or forwards and backwards propagation.

In the long paper [48] we deal in detail with the point source case, that is, the propagation of an acoustic pulse generated by a point source over a layered random medium. Here I will describe only the reflection of acoustic plane wave pulses.

4.1 Pulse reflection from randomly layered media

The acoustic pressure $p(t,z)$ and velocity $u(t,z)$ satisfy the continuity and momentum equations

$$\begin{aligned} \frac{1}{K} p_t + u_z &= 0 \\ \rho u_t + p_z &= 0 \end{aligned} \tag{19}$$

Here ρ is the material density and K the bulk modulus. As in [48] we assume for simplicity that the density has no random variation

$$\rho(z) = \begin{cases} \rho_0, & z > 0, \\ \rho_1, & z < 0 \end{cases} \tag{20}$$

with ρ_0 and ρ_1 constants. For the bulk modulus we assume that

$$K^{-1}(z) = \begin{cases} K_0^{-1}, & z > 0, \\ K_1^{-1}(z)\left(1 + \nu(\frac{z}{\varepsilon^2})\right), & z < 0 \end{cases} \tag{21}$$

with K_0 a constant, $K_1(z)$ a smooth deterministic function of z and $\nu(s)$ a bounded stationary random function with mean zero, representing the fluctuations in K^{-1}. Note that they vary on the scale ϵ^2, where ϵ is a small parameter. If z is measured in kilometers and the fluctuations vary on the scale of a few meters then a value of ϵ around 0.05 captures the scale separation we wish to model. We assume that the random function $\nu(s)$ has a correlation length of order one so that the correlation length of $\nu(z/\epsilon^2)$ is of order ϵ^2 in kilometers (about 2.5 meters for $\epsilon = 0.05$). The mean sound speed c is given by

$$c(z) = \begin{cases} c_0 = \sqrt{\frac{K_0}{\rho_0}}, & z > 0 \\ \bar{c}(z) = \sqrt{\frac{K_1(z)}{\rho_1}}, & z < 0 \end{cases} \tag{22}$$

Note that the fluctuations in the sound speed are not assumed to be small. The estimation of the vertical correlation length of the inhomogeneities in the lithosphere from well-log data is considered in [51]. They found that 2-3 m is a reasonable estimate of the correlation length of the fluctuations in sound speed.

For $t < 0$ a normally incident plane wave solution in $z > 0$ has the form

$$\begin{aligned} u(t,z) &= \frac{1}{\sqrt{\epsilon}} \frac{1}{\sqrt{\rho_0 c_0}} f\left(\frac{t + z/c_0}{\epsilon}\right) \\ p(t,z) &= -\frac{1}{\sqrt{\epsilon}} \sqrt{\rho_0 c_0} f\left(\frac{t + z/c_0}{\epsilon}\right) \end{aligned} \tag{23}$$

Here f is the pulse shape function which is assumed to vanish for negative arguments and to have support that is of order one in the macroscopic t units that are seconds. With $\epsilon = 0.05$, the pulse width is about 50 msec or, with a speed of 3 km/sec, 150 meters. The multiplicative factor $1/\sqrt{\epsilon}$ in (23) makes the total energy of the incident plane wave pulse independent of ϵ. Continuity of p and u at the interface $z = 0$ makes (19) and (23) a complete problem. We are interested in $p(t,0)$ or $u(t,0)$ for $t > 0$, the pressure or velocity at the interface, and this involves the solution of a complicated random scattering problem because of the form (21) of K^{-1}.

4.2 SCALE SEPARATION

The scaling that we have chosen, and the asymptotic limit $\varepsilon \to 0$ that we will consider, models well problems in reflection seismology and is quite different from transport theoretic scaling. The main differences are that the fluctuations are not assumed to be small and the typical wavelength of the probing pulse (150 m) is small with respect to the probing depth (5 km, say) but large compared to the correlation length (2-3 m). The parameter ε is then the ratio of the (typical) wave length to propagation depth, as well as the ratio of correlation length to wave length. This is a particularly interesting scaling limit mathematically because it is a high frequency limit with respect to the large scale variations of the medium that we want to detect, but it is a low frequency limit with respect to the fluctuations, whose effect acquires a canonical form independent of details.

Is this model realistic and can it be used effectively? One argument that can be made against it is this: There is no real scale separation in sound speed fluctuations, as one can see from well-logs [52], so this neat way of dealing with fluctuations, background and probing pulse cannot possibly be right, even if it can handle large fluctuations. Another is that perfectly layered random media are an unacceptable idealization.

Regarding scale separation, it is fair to say that the scope of the analytical theory that has been developed, and is described briefly here, is well beyond anything that could be expected from any theory that deals with strong fluctuations in a serious way. Radiative transport theory is more robust because the fluctuations are assumed to be small, and then it is not necessary to have scale separation (correlation lengths and wave lengths are comparable). Moreover, the analytical tools that emerge from the asymptotic scale separation theory are far more flexible and robust than the crude thinking with scales implies. Discontinuities and imperfections that are comparable to the pulse width can be handled by the theory and do not make it unusable. The problem is that the theory is not easy to follow, it is analytically difficult to implement and not nearly enough has been done to test it in situations that push against the scale separation assumptions. The statistical analysis of well-log data that was done in [51], that produced the estimate of 2-3 m for the correlation length of the sound speed fluctuations, is quite thorough, but perhaps more can be done here also.

The modifications to the theory that are needed to account for imperfect layering are far more important than anything missed by scale separation asymptotics. This goes back to the localization-delocalization transition that I have mentioned several times already. It remains a big gap in our understanding of wave propagation in random media.

4.3 Localization regime asymptotics

We will consider the reflected pressure $p_{refl}(t,0)$, at $z=0$ and $t>0$, which is the total pressure minus the incident pressure (23). After a time of order ϵ, the duration of the incident pulse, the two are the same. Of particular interest is the two-time reflected average pressure intensity.

$$I(t,\bar{t}) = \frac{1}{\rho_0 c_0} < p_{refl}(t+\frac{\epsilon\bar{t}}{2},0)p_{refl}(t-\frac{\epsilon\bar{t}}{2},0) > \tag{24}$$

with the angular brackets denoting statistical average. The factor $1/\rho_0 c_0$ is a normalization.

For simplicity, we will assume in the sequel that there is no macroscopic discontinuity at $z=0$ so that $\rho_0=\rho_1$ and $K_0^{-1}=K_1^{-1}(0)$.

Note that the time offset in (24) is proportional to the pulse width ϵ. The reason for this is that for time offsets of more than a few pulse widths the reflected signals are essentially uncorrelated. Moreover, in the absence of discontinuities in the medium, $< p_{refl}(t, 0) >$ is essentially zero except for a time of order ϵ near $t = 0$ when the reflection from the interface $z = 0$ is felt.
That is, there is no coherent backscattering. We formulated a scattering problem where the quantity of interest is as directly related to the medium fluctuations as is possible.

Fix a $t > 0$, not close to zero, and a small ϵ. Since $I(t, \bar{t})$ is essentially zero for large $\bar{t}$ we can introduce its (essentially local) Fourier transform

$$\Lambda(t,\omega)|\hat{f}(\omega)|^2 = \int e^{i\omega\bar{t}} I(t,\bar{t}) d\bar{t} \tag{25}$$

in which Λ is the normalized local power spectral density. The normalization is $|\hat{f}(\omega)|^2$ with $\hat{f}(\omega)$ the Fourier transform of the pulse shape function $f(t)$. The two-time intensity function can be written as

$$I(t,\bar{t}) = \frac{1}{2\pi} \int |\hat{f}(\omega)|^2 \Lambda(t,\omega)\ e^{-i\omega\bar{t}} d\omega \tag{26}$$

The main thrust of our theoretical work in [45, 48, 46] is that in the limit $\epsilon \to 0$ the local power spectral density can be calculated by solving a system of partial differential equations where

$$\Lambda(t,\omega) = W^1(0,t,\omega) \tag{27}$$

and the $W^N(z,t,\omega),\ N \geq 0$ satisfy the equations

$$\frac{\partial W^N}{\partial z} + \frac{2N}{\bar{c}(z)} \frac{\partial W^N}{\partial t} - \frac{2\alpha\omega^2 N^2}{\bar{c}^2(z)} \left\{ W^{N+1} - 2W^N + W^{N-1} \right\} = 0 \tag{28}$$

for $-L < z \leq 0$, with

$$W^N(-L,t,\omega) = \delta(t)\delta_{N,0} \tag{29}$$

Here the mean sound speed $\bar{c}(z)$ is given by (22) and $\alpha > 0$ is the noise intensity level of the fluctuations

$$\alpha = \frac{1}{4} \int_0^\infty < \nu(s)\nu(0) > ds \tag{30}$$

The length L is arbitrary, provided that for any given $t > 0$ for which we want to calculate $\Lambda(t,\omega)$ it satisfies

$$L > c_{max} \frac{t}{2} \tag{31}$$

with c_{max} the maximum speed $\bar{c}(z)$ in $z \leq 0$. Because of the hyperbolic nature of the equations (28) it is easy to see (and explained in the references) that the choice of L satisfying (31) does not affect $\Lambda(t,\omega)$ given by (27).

4.4 Time domain localization

There is no quick and simple way to explain the result (27)-(31) that relates the local power spectral density $\Lambda(t,\omega;\bar{c}(.))$, the mean sound speed profile $\bar{c}(z)$ and the noise intensity level α. But we will now make several remarks that will help explain the nature of this relationship.

From the definition (24) and (25) it is clear that $\Lambda(t,\omega)$ is a local Fourier transform but it is not necessarily positive as it would have to be if $p_{refl}(t,0)$, the reflected pressure, were a stationary process in t so that $I(t,\bar{t})$ were independent of t. However, in the limit $\epsilon \to 0$, and hence when ϵ is small, $\Lambda(t,\omega)$ given by (27)-(31) is indeed positive. For a general profile $\bar{c}(z)$ it cannot be computed explicitly but for $\bar{c}(z)=\bar{c}$, a constant, it has the form

$$\Lambda(t,\omega) = \frac{\frac{\alpha\omega^2}{\bar{c}}}{\left(1+\frac{\alpha\omega^2}{\bar{c}}t\right)^2} \tag{32}$$

In terms of the localization length [49] at frequency ω

$$l(\omega) = \frac{\bar{c}^2}{2\alpha\omega^2} \tag{33}$$

we can write (32) in the form

$$\Lambda(t,\omega) = \frac{1}{2}\frac{\bar{c}l(\omega)}{\left(l(\omega)+\frac{\bar{c}t}{2}\right)^2} \tag{34}$$

As shown in [49] and the many references cited there, the localization length at frequency ω is a measure of the depth of penetration of a time harmonic wave with this frequency into a randomly layered medium with uniform sound speed $\bar{c}$ and noise level α for the fluctuations. Wave energy does not penetrate much below this length. If $T(L,\omega)$ is the time harmonic transmission coefficient for a randomly layered medium of width L, with ω the frequency of the incident plane wave, then

$$\lim_{L\to\infty}\frac{1}{L}\log|T(L,\omega)| = \frac{-1}{l(\omega)}$$

with probability one. This defines the localization length $l(\omega)>0$, which is always positive for a large class of random media. It cannot be computed explicitly but in the low frequency limit it has the form (33). The lower the frequency the deeper the penetration of the waves into the randomly layered medium.

In the time domain, the normalized local power spectral density of the reflected signal at a fixed time t, $\Lambda(t,\omega)$ in (34), has a maximum $\omega_{max}=\omega_{max}(t)$ that depends on time. From (34) the maximum is calculated to be

$$\omega_{max} = \sqrt{\frac{\bar{c}}{\alpha t}} \tag{35}$$

In a more physical way [49] we can say that the maximum of the local power spectral density at time t occurs for that $\omega=\omega_{max}$ for which

$$l(\omega_{max}) = \frac{\bar{c}t}{2} \tag{36}$$

Thus, for the frequency for which the localization length equals the mean distance traveled into the medium, we have the maximum contribution to the noise spectrum of the reflected signal. This is a stochastic resonance relation that identifies precisely the main source of noise in the reflected signals.

It is because of wave localization and its manifestations in the time domain described above that signals reflected by randomly layered media are so noisy. From (32) we find, by integrating over ω, that the envelope of the root mean square of the reflected pulse is of the form constant $\times\, t^{-3/4}$. Thus, the fluctuations in the reflected signal decay very slowly, indicating that a great deal of multiple scattering is taking place and that wave localization is dominant.

We can interpret (28) as a hierarchy of equations for moments associated with the scattering problem [48]. The infinite hierarchy ($N \geq 0$ in (28)) indicates that the second moment that we are interested in (I of (24) or Λ of (25)) cannot be computed separately from all higher moments (the $W^N,\ N > 2$ in (28)). This is another manifestation of localization.

When we use the parameters of section 4.2 that are typical in reflection seismology we find that the minimum localization length occurs in the 20-30 Hz regime and is about 15-20 km [51]. This means that random inhomogeneities will effectively prevent probing below this depth because all the wave energy is reflected to the surface by multiple scattering.

What is missing at present is a more general theory that allows us to compute the changes in the one dimensional theory that occur when small three dimensional inhomogeneities are introduced into the model. We need a more general theory that lets us go from localization to transport as the random layering is reduced and isotropic inhomogeneities replace it.

4.5 Statistical inverse problems

I will describe briefly how the mean sound speed profile $\bar{c}(z)$ can be estimated from observations of $p_{refl}(t,0)$ or

$$R_f(t) = \frac{1}{\sqrt{\rho_0 c_0}} p_{refl}(t,0) \tag{37}$$

in which dependence of the pulse shape function f is indicated. The inversion strategy is based on one more fact about the reflected signal $R_f(t)$, in addition to (27)-(31). It is that as ϵ tends to zero $R_f(t)$ becomes approximately a Gaussian process. It has not been possible to prove this so far but there are some good heuristic indications that it is true [45] and extensive numerical simulations corroborate it very well [48]. From the Gaussian property of $R_f(t)$ we conclude that

$$\frac{1}{|\hat{f}(\omega)|^2} \int e^{i\omega\bar{t}} R_f\left(t + \frac{\epsilon\bar{t}}{2}\right) R_f\left(t - \frac{\epsilon\bar{t}}{2}\right) d\bar{t} = \hat{\Lambda}(t,\omega) \tag{38}$$

is approximately, when ϵ is small, an exponential random variable with mean $\Lambda(t,\omega)$ given by (27)-(31), when $\bar{c}(z)$ is known. Moreover, for distinct $0 < t_1 < t_2 < ... < t_{N_t}$ and $0 < \omega_1 < \omega_2 < ... < \omega_{N_f}$, where N_t and N_f are integers, the random variables $\left\{\hat{\Lambda}(t_j,\omega_l)\right\}$ are independent with exponential distribution having mean $\{\Lambda(t_j,\omega_l)\}$.

The inversion strategy is now this: Depending on the available data, fix a set of time points $\{t_j\}$ and frequencies points $\{\omega_l\}$ as above. For each realization of $R_f(t)$ that is available, estimate $\hat{\Lambda}(t_j,\omega_l)$ from (38). This is actually a very delicate step that must be done carefully as we discuss in [48], Appendix E. Then form

$$O(\bar{c}) = \prod_{realiz}\prod_{j=1}^{N_t}\prod_{l=1}^{N_f}\frac{e^{-\hat{\Lambda}(t_j,\omega_l)/\Lambda(t_j,\omega_l;\bar{c}(.))}}{\Lambda(t_j,\omega_l;\bar{c}(.))} \tag{39}$$

where the first product is over different independent realizations. This is the likelihood functional for the estimates $\hat{\Lambda}$, given a known mean speed profile $\bar{c}(z)$. We now choose $\bar{c}(z)$ in order to maximize this functional. This is a rather usual maximum likelihood estimation *except* that now the maximization must be done over the profiles $\bar{c}(.)$ which in turn determine $\Lambda(t,\omega;\bar{c}(.))$ in (39) via the partial differential equations (28)-(29) and the relation (27).

The most convenient way to solve the maximization problem for (39), and thus estimate $\bar{c}(z)$, is to assume that it is piece-wise linear over a few macroscopically large layers and then maximize O over a finite set of speeds $\bar{c}^1,\bar{c}^2,...,\bar{c}^{N_z}$. These speeds are approximations of $\bar{c}(z)$ at successively larger depths numbered from 1 to N_z. Moreover, because of the hyperbolic nature of (28)-(29) the maximization can be done one layer at a time with increasing depth. This avoids the difficult problem of finding the maximum of a complicated function of several variables. Physically this layer peeling process makes sense because there is a direct relation between the sound speed profile up to a certain depth and the smallest time before which the rest of the medium is not felt in the reflected signal $R_f(t)$.

Of course we need a lot of independent realizations to get reasonable results and this is unrealistic in a geophysical context. But it is important in principle to make this strategy work and amazingly enough it does [50], [48]). It is amazing because we are trying to determine the smooth, mean speed profile from the reflected signals that are swamped by fluctuations due to multiple scattering. The computational and other implementation details are described in [48]).

Could we do this kind of inversion from extremely noisy reflections if we only had one realization? Yes, if we have reflection measurements at different offsets (distances from the source) on the interface, generated by a point source over a randomly layered medium [53]. This is a very difficult problem that requires a great deal of numerical computation. The inversion is not as good as in the plane wave case (with many realizations) but it is reasonably good and, in any case, it shows that the strategy does work. But improving the results requires very careful attention to a host of implementation issues that can be settled only empirically, by trial and adjustment, at present.

An interesting discussion of reflections from time reversed reflections, their statistical properties and their relation in turn to the hierarchy of moments equations (28) is given by Clouet and Fouque [54]. This work should have important applications in statistical inverse problems of geophysical interest.

Another application of time domain localization asymptotics is to surface water waves over a rough bottom [55].

4.6 Reflection and transmission of time harmonic plane elastic waves

We have described a variety of results for *acoustic* pulse reflection from randomly layered media, emphasizing time domain effects. For geophysical applications we must also consider *elastic* waves in randomly layered media. The analytical difficulties in extending the theory that we briefly described above to the elastic case are enormous, mainly because there are two wave modes, P and S waves, that are coupled by the inhomogeneities. In [56] we extended the scale-separation asymptotic theory to time harmonic, obliquely incident elastic plane waves. We calculate in detail mode coupling in reflection and transmission, with various kinds of interfacial discontinuities. It is surprising that so many things can be calculated analytically and in such detail, given the complexity of the problem.

However, despite considerable efforts we have not been able to extend the results to the time domain. The hierarchy of moment equations that we used in the analysis of acoustic pulse reflection does not seem to work for elastic wave pulse reflection. The analysis of reflections for elastic wave pulses generated by a point source, the analog of the analysis carried out in [48] for acoustic waves, seems to be out of reach at present.

4.7 Pulse stabilization and imaging

We have focused mostly on reflection in the time domain because the bulk of the measurements that can be made in geophysics, in nondestructive testing with ultrasound and elsewhere are surface measurements. However, transmission is also important as is the analysis of reflections from imbedded discontinuities in a randomly layered medium. The vicinity of the front of the pulse, or the vicinity of first arrival from the discontinuity, has an interesting structure that can be analyzed in considerable detail. This is called the O'Doherty-Anstey theory because it was first discussed by these two geophysicists in the early seventies [57]. The main point is that if the fluctuations are weak and the pulse is followed with its random speed, then it will appear to stabilize (not fluctuate) and become broader as it advances into the medium. This is discussed in detail in [48] where many other papers are cited.

What if the fluctuations are not weak, and we have scale separation as described above? Do we have an O'Doherty-Anstey theory? This question was answered in [58, 59] by overcoming what was the main obstacle before: finding the right random speed with which to center the advancing pulse. The fact that the advancing pulse spreads and loses energy (to fluctuations in its coda) is not so surprising and is true for general random media, not only layered media, although the fluctuations must be weak. What is surprising, and not generally known or anticipated in the geophysics literature, is that in the case of large fluctuations the centering speed is not the local random speed but a function of it, and the centered pulse stabilizes with probability one with minimal spreading (relative to other centerings).

In [60], Solna shows how this theory can be used to improve the resolution of discontinuity identification in a random medium. He also extends the O'Doherty-Anstey theory to a class of locally layered random media, that is, he allows for slow horizontal variations.

References

[1] P.W. Anderson, *Absence of diffusion in certain random lattices*, Phys. Rev., 109 (1958), pp. 1492-1505.

[2] J.B. Keller, *Stochastic Equations and Wave Propagation in Random Media*, Proc. Symp. Appl. Math., 16, 145–170, Am. Math. Soc., McGraw–Hill, New York, 1964.

[3] A.Ishimaru, *Wave propagation and scattering in random media*, vol. I,II, Academic Press, New York, 1978.

[4] K.Aki and R.S.Wu, Multiple scattering and energy transfer of seismic waves – separation of scattering effect from intrinsic attenuation –II. Application of the theory to Hindu Kush region, in *Seismic wave scattering and attenuation*, vol. I (R.S.Wu and K.Aki, eds), pp. 49-80, 1988.

[5] M.C. Fehler and H. Sato, *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*, AIP Press, 1998.

[6] P. Sheng, *Introduction to wave scattering, localization, and mesoscopic phenomena*, Academic Press, 1995.

[7] A. Figotin and L.A. Pastur, *Spectra of Random and Almost-Periodic Operators*, Springer-Verlag, 1992.

[8] R. Knapp, G. Papanicolaou and B.S. White, *Transmission of waves by a nonlinear random medium*, J. Stat. Physics, 63, pp. 567-584, 1991.

[9] J.C. Bronski, *Nonlinear Wave Propagation in a Disordered Medium* J. Stat. Phys., 92, 5/6, 1998.

[10] J.C. Bronski *Nonlinear Scattering and the Analyticity Properties of Solitons*, Journal of Nonlinear Science, 8, pp. 161-182, 1998.

[11] J.C. Bronski, D.W. McLaughlin and M.J. Shelley, *Stability of Time-Harmonic Localized Solutions in a One Dimensional Disordered Nonlinear Medium*, J. Stat. Phys., 88, 1997.

[12] J. Garnier, *Transmission of solitons through random media*, to appear in SIAM, J. Appl. Math.

[13] J. Garnier, PhD Thesis, Ecole Polytechnique, 1996.

[14] J.F. Claerbout, *Imaging the Earth's Interior*, Blackwell, Boston, 1985.

[15] N. Bleistein, J.K. Cohen and J.W. Stockwell Jr., *Mathematics of Multidimensional Seismic Inversion*, pp. 1-260, Colorado School of Mines, 1998.

[16] W.W. Symes, *Mathematics of Reflection Seismology*, pp. 1-85, Rice University, 1995.

[17] L.Ryzhik, G.Papanicolaou and J.Keller, Transport Equations for Elastic and Other Waves in Random Media, Wave Motion 24, pp. 327-370, 1996.

[18] M.C.W. van Rossum and Th.M.Nieuwenhuizen, Multiple scattering of classical waves: from microscopy to mesoscopy and diffusion, to appear in Phys. Rev. E.

[19] P.Gérard, P.Markowich, N.Mauser and F.Poupaud, Homogenization limits and Wigner transforms, Comm. Pure Appl. Math., 50, pp. 323-380, 1996.

[20] L.Ryzhik, G.Papanicolaou and J.Keller, Transport equations in a half space, Comm. PDE's, 22, pp. 1869-1911, 1997.

[21] P.Stott, A transport equation for the multiple scattering of electromagnetic waves by a turbulent plasma, Jour. Phys. A, 1, pp. 675-689, 1968.

[22] K.Watson and J.L.Peacher, Doppler shift in frequency in the transport of electromagnetic waves in an underdense plasma, Jour. Math. Phys., 11, pp. 1496-1504, 1970.

[23] K.Watson, Multiple scattering of electromagnetic waves in an underdense plasma, Jour. Math. Phys., 10, pp. 688-702, 1969.

[24] C.W.Law and K.Watson, Radiation transport along curved ray paths, Jour. Math. Phys., 11, pp. 3125-3137, 1970.

[25] Yu.Barabanenkov, A.Vinogradov, Yu.Kravtsov and V.Tatarskii, Application of the theory of multiple scattering of waves to the derivation of the radiative transfer equation for a statistically inhomogeneous medium, Radiofizika, 15, pp. 1852-1860, 1972. English translation pp. 1420-1425.

[26] I.M.Besieris and F.D.Tappert, Propagation of frequency modulated pulses in a randomly stratified plasma, Jour. Math. Phys., 14, pp. 704-707, 1973.

[27] M.S.Howe, On the kinetic theory of wave propagation in random media, Phil. Trans. Roy. Soc. Lond., 274, pp. 523-549, 1973.

[28] I.M. Besieris, W. Kohler and H. Freese, A transport-theoretic analysis of pulse propagation through ocean sediments, J. Acoust. Soc. Am., 72, pp. 937-946, 1982.

[29] Yu.Barabanenkov, Yu.Kravtsov, V.Ozrin and A.Saichev, Enhanced backscattering in optics, Progress in Optics, 29, pp. 67-190, 1991.

[30] P. Martin and G. Emch, A rigorous model sustaining van Hove's phenomenon, Helv. Phys. Acta, 48, p. 59, 1967.

[31] H. Spohn, Derivation of the transport equation for electrons moving through random impurities, Jour. Stat. Phys., 17, pp. 385-412, 1977.

[32] G.Dell'Antonio, Large time small coupling behavior of a quantum particle in a random potential, Annals of Inst. H. Poincare, section A, 39, p. 339, 1983.

[33] T. Ho, L. Landau and A. Wilkins, On the weak coupling limit for a Fermi gas in a random potential, Rev. Math. Phys., 5, pp. 209-298, 1993.

[34] L. Erdos and H.T. Yau, preprint, 1997.

[35] A.Fannjiang, unpublished notes, 1991.

[36] S.Chandrasekhar, *Radiative transfer*, Dover, New York, 1960.

[37] G.Papanicolaou, L.Ryzhik and J.Keller, *On the stability of the P to S energy ratio in the diffusive regime*, Bulletin of the Seismological Society of America, 86, pp. 1107-1115, 1996.

[38] J.A. DeSanto and G.S. Brown, *Analytical techniques for multiple scattering from rough surfaces*, in Progress in Optics XXIII (E. Wolf, ed.), pp. 1-62, Elsevier, Amsterdam, 1986.

[39] A.Voronovich, *Wave Scattering from Rough Surfaces*, Springer Verlag, 1991.

[40] G. Bal, J.B. Keller, G. Papanicolaou and L. Ryzhik, *Transport Theory for Acoustic Waves with Reflection and Transmission at Interfaces*, preprint, 1998.

[41] R.L.Weaver, *On diffuse waves in solid media*, Jour.Acoust.Soc.Am., 71, pp. 1608-1609, 1982.

[42] G. Bal, PhD Thesis, University of Paris VI, 1997.

[43] O. Dorn, *Diffuse tomography*, PhD Thesis, University of Münster, 1997.

[44] L.Brekhovskikh, *Waves in Layered Media*, Academic Press ,1980.

[45] R. Burridge, G. Papanicolaou, P. Sheng, and B. White, *Probing a random medium with a pulse*, SIAM J. Appl. Math., 49, pp. 582-607, 1989.

[46] G. Papanicolaou and S. Weinryb, *A functional limit theorem for waves reflected by a random medium*, Applied Mathematics and Optimization, 30, pp. 307-334, 1994.

[47] W. Kohler and G. Papanicolaou, *Power statistics for wave propagation in one dimension and comparison with radiative transport theory I,II*, J. Math. Phys., 14, pp. 1733-1745, 1973, and 15, pp. 2186-2197, 1974.

[48] M. Asch, W. Kohler, G. Papanicolaou, M. Postel and B. White, *Frequency Content of Randomly Scattered Signals*, SIAM Review, 33, pp. 526-629, 1991.

[49] P. Sheng, B. White, Z.Q. Zhang and G.Papanicolaou, *Wave localization and multiple scattering in randomly-layered media*, in 'Scattering and Localization of Classical Waves in Random Media,' Ping Sheng, editor, World Scientific, pp. 563-619, 1990.

[50] B. White, P. Sheng, M. Postel, and G. Papanicolaou, *Probing through cloudiness: Theory of statistical inversion for multiply scattered data*, Phys. Rev. Lett., 63, pp. 2228-2231, 1989.

[51] B. White, P. Sheng, and B. Nair, *Localization and backscattering spectrum of seismic waves in stratified lithology*, Geophysics, 55, pp. 1158-1165, 1990.

[52] F. Herrmann, *A scaling medium representation, a discussion of well-logs, fractals and waves*, PhD Thesis, Delft University, 1997.

[53] M. Asch, W. Kohler, G. Papanicolaou, M. Postel and B. White, *Statistical inversion from reflections of spherical waves by a randomly layered medium*, Waves in Random Media, 6, pp. 293-334, 1996.

[54] J.F. Clouet and J.P. Fouque, *A time-reversal method for an acoustical pulse propagating in randomly layered media*, Wave Motion, 25, pp. 361-368, 1997.

[55] A. Nachbin and G. Papanicolaou, *Water waves in shallow channels of rapidly varying depth*, JFM, 241, pp. 311-332, 1992.

[56] W. Kohler, G. Papanicolaou and B. White, *Localization and mode conversion for elastic waves in randomly layered media I,II*, Wave Motion, 23, pp. 1-22 and 181-201, 1996.

[57] R.F. O'Doherty and N.A. Anstey, *Reflections on amplitudes*, Geophysical Prospecting, 19, pp. 430-458, 1971.

[58] J.F. Clouet and J.P. Fouque, *Spreading of a pulse traveling in random media*, Annals of Applied Probability, 4, pp. 1083-1097, 1994.

[59] R. Burridge, P. Lewicki and G. Papanicolaou, *Pulse stabilization in a strongly heterogeneous layered medium*, Wave Motion, 20, pp. 177-195, 1994.

[60] K. Solna, PhD Thesis, Stanford University, 1997.

George Papanicolaou
Department of Mathematics
Stanford University
Stanford CA 94305
papanico@math.stanford.edu
URL: http://georgep.stanford.edu