Sifting Noise: The Role of Probability in Imaging[1]

Josselin Garnier[2]

Prisme N°25

November 2012

[1] This text was translated by Richard Crabtree. The original French version is the transcription of Josselin Garnier's presentation given at the Cournot Centre's seminar series, "The Probabilism Sessions" in Paris. The video (in French) is available on the Cournot Centre's website: www.centre-cournot.org.

[2] Josselin Garnier is Professor of mathematics at the University of Paris VII and a researcher at the LPMA (Probabilities and Random Models Laboratory) and the Jacques-Louis Lions Laboratory in Paris. He is a junior member of the *Institut Universitaire de France*. Professor Garnier received the Blaise Pascal Prize from the French Academy of Sciences in 2007 and the Felix Klein Prize from the European Mathematical Society in 2008.

© Cournot Centre, November 2012

Summary

Imaging techniques using waves to probe unknown media have long existed. Classically, these techniques can be divided into a phase of data gathering and a phase of data processing. During the data-gathering phase, waves are emitted by a source or source array, propagated through the medium being studied, and are then recorded by a receiver array. The processing phase consists in extracting information about the medium from the data recorded by the receivers. Recently, new ideas have emerged driven by observations made during time-reversal experiments. Based on these observations, new imaging methods have been developed using cross correlations of the signals recorded by sensor arrays. Mathematical analysis has shown that the cross correlation of signals recorded by two passive sensors essentially contains as much information about the medium as the signal that would have been recorded if one of the sensors were active (emitter) and the other passive (receiver). The important point demonstrated by this analysis is that uncontrolled sources of ambient noise can be used instead of controlled sources to compute cross correlations and use them for imaging. This possibility has attracted the attention of researchers in mathematics, in the domain of probabilities, for profound theoretical reasons, because the idea of useful noise overturns the customary distinction between signal and noise. This has also been the case in seismology for obvious practical reasons concerning the sparsity of sources (earthquakes) and the impossibility of controlling them. The aim of this paper is to describe how the idea of exploiting ambient noise to address problems of imaging took shape.

© Cournot Centre, November 2012

This text is based on the research I have been conducting over the last few years, mainly with my colleague George Papanicolaou of Stanford University. When presenting our work to audiences of mathematicians, we usually start by explaining the methods and results and then give the proofs. The present text will rather retrace the story that led us to study noise with a view to applications in the field of imaging. Our goal was to build images from signals hitherto considered to be pure noise, that is to say completely unusable signals.

According to one definition, noise corresponds to random, unwanted or even parasitic signals overlying some useful signals. This definition presupposes a distinction between useful and non-useful signals, and all noise is rejected in the second category. While working with seismologists, we realized that what they called "noise", and considered to be useless, had never been subject to careful study, and consequently had never had the chance to be considered useful. Thus, for nearly a century, seismologists have been discarding signals that actually contain a wealth of information.

The context for this story is imaging, the aim of which is to produce pictures of an object in the broadest sense – the human body, for example, in the case of medical imaging, or the Earth's crust in the case of seismology. The general principle consists in using waves to probe the medium in question. The waves are emitted by a source or a source array, they propagate through and interact with the medium, and they are then recorded by a sensor array. These recordings constitute the data set that is then processed to extract the desired information (position, shape, and characteristics of a buried object, for example). Imaging may involve reflection techniques, as in optics, for example, where light is projected onto the object and whatever it reflects is photographed, or in ultrasound echography, where ultrasound waves are reflected by the obstacles they encounter in the human body. Other techniques involve transmission, for example X-ray radiography. The classic contribution of probability consists in quantifying the impact of noise and limiting it as far as possible. From the probabilistic perspective, the important thing for imaging specialists is to eliminate noise, in other words, to increase the signal-to-noise ratio. That is their main objective; ours is very different.

Let us consider the context is geophysics, where the aim is to produce images of the interior of the Earth. Applied geophysics focuses above all on the upper layers

© Cournot Centre, November 2012

of the Earth's crust in the search for mineral, oil and gas deposits. Seismology is a subdiscipline of geophysics. It studies the propagation of seismic waves. By constructing a map of the propagation velocity of seismic waves in the Earth's crust, we can obtain a picture of the Earth's crust, because the propagation velocity is characteristic of the medium. The classic method of constructing such maps, used since the 19th century, consists in analysing the signals from a seismometer network spread over the Earth's surface. The seismic waves produced by earthquakes are signals that scientists seek to record and analyse.

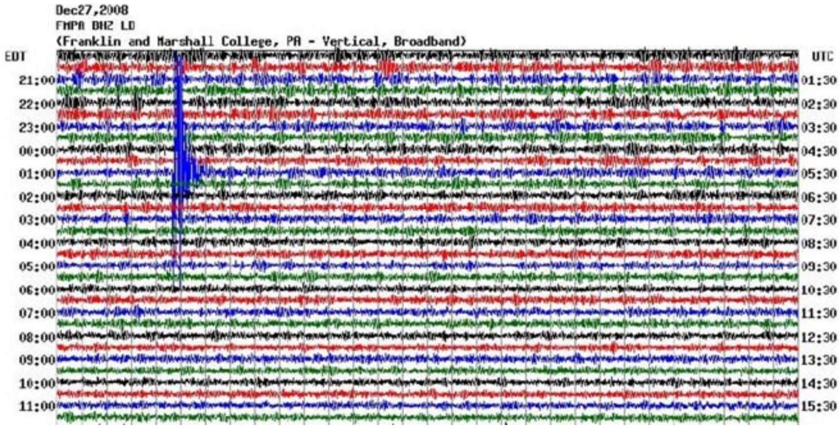


Figure 1: A seismogram

In this seismogram (Figure 1), signals were recorded over a period of 14 hours, with each line corresponding to a half hour of recording. Essentially, the only thing that is recorded is noise, noise and yet more noise, but at a certain moment, the passage of a seismic wave produced by an earthquake is recorded. Seismologists are interested in the little section of recording that corresponds to the passage of the seismic wave. The only information they use is the arrival time; they record the time at which the seismic wave passes through the site of the seismometer.[3] By recording

[3] In fact, seismic waves are elastic and an earthquake can emit several different types of seismic waves, which propagate at different speeds. Several arrival times can thus be detected, each corresponding to a particular type of wave.

© Cournot Centre, November 2012

the arrival times of the seismic wave over a seismometer array, it is possible to make an estimation of the velocity map of the propagation of seismic waves through the Earth's crust. If the positions of the seismometers have been clearly identified, and the arrival times of the seismic wave have been accurately recorded, then it is possible to estimate the wave propagation velocity, by means of least-square rules.[4]

Seismologists also often adopt a Bayesian approach. Figure 2 illustrates the underlying principle. An *a priori* model of a seismic velocity map is defined (it could even be a uniform *a priori* model, since our first intuition is that wave velocity will be uniform). An earthquake happens somewhere, the arrival times of the seismic wave created by the earthquake are recorded, and then, according to how these arrival times fit or diverge from the predictions, the *a priori* model is updated to produce an *a posteriori* model. So that is a rough outline of classic seismology.

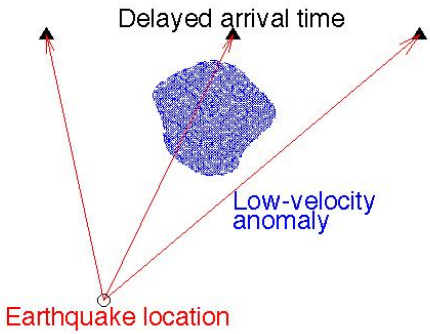


Figure 2: Classic seismology model

It produces the type of map presented in Figure 3, which shows the variations in seismic wave velocity on large scale, at a global level. Here we see the variations in depth following a line running from Europe to Japan. These maps exist on a large scale or on much more reduced scales.

[4] Actually these are slightly more complicated rules of travel time tomography.

© Cournot Centre, November 2012

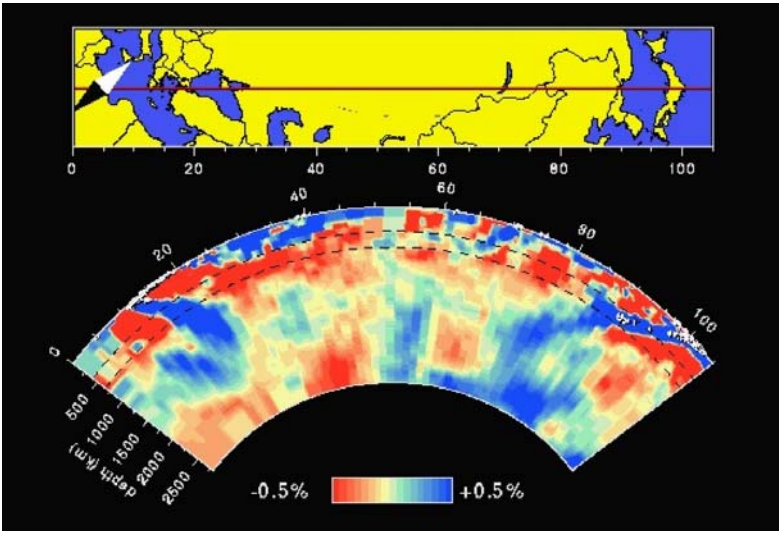


Figure 3: Maps of variations in seismic wave velocity on a large scale from Europe to Japan

So classic seismology uses the information contained in the arrival times of seismic waves. It operated in this way for nearly a century. In the 1950s, however, (with the work of Keiiti Aki [Aki and Chouet, 1975]), interest was also focused on the information contained in the "coda" (the small oscillations following the wave front), where the rate of decrease provides information about the diffusion rate of seismic waves in the medium. By extracting the attenuation coefficient, one can obtain information about the properties of the medium. In the 1980s, researchers started trying to exploit the coda. Whether working on arrival time tomography or coda analysis, however, one has to wait for an earthquake powerful enough for the seismometers to record a seismic wave. This is easy in some regions of the world, like California, but less so in others where the sources (earthquakes) are infrequent and it is difficult to do imaging.

Nevertheless, as we can see in the seismogram in Figure 1, outside the brief interval corresponding to the passage of a seismic wave created by an earthquake, seismic background noise can be observed.

© Cournot Centre, November 2012

Now I would like to explain how the idea of studying and exploiting seismic noise arose. At this stage, the story gets rather complicated; it follows a circuitous path originating not far from the rue Gazan in Paris, at the *École Supérieure de Physique et de Chimie Industrielles* (School of Industrial Physics and Chemistry). Physicists – acousticians from the Langevin Institute – developed a new technique for medical applications called the "time reversal of ultrasonic waves".

This technique is based on a mathematical property of the wave equation known as time-reversal invariance.[5] This gave Mathias Fink the idea for an original experiment. The time-reversal experiment is based on a special apparatus called a time-reversal mirror, consisting of a collection of transducers – little devices that have two operating modes and can be used sometimes as transmitters and other times as receivers. In a time-reversal experiment, this mirror is first used as a receiver array and then as a transmitter array. A time-reversal experiment proceeds as follows (Figure 4). A source buried in the medium emits an acoustic pulse.

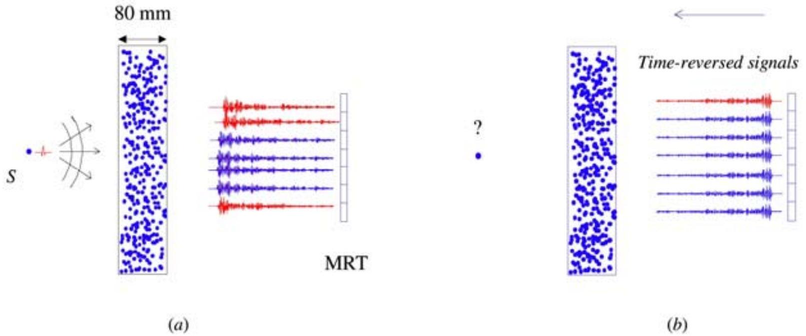


Figure 4: Time reversal – the principle underlying the experiment (Tourin et al., 2000)

The pulse then propagates through the medium, which may or may not be homogeneous, and the signals arriving at the time-reversal mirror (TRM in Figure 4) are recorded. So the first part of the experiment is a phase of data capture: the signals emitted by a source and recorded by the mirror are stored on a computer. The

[5] If a function is a solution to the wave equation, and therefore a function of time and space, then the same function with the time reversed is also a solution to the wave equation.

© Cournot Centre, November 2012

signals are time-reversed, so that what came first now comes at the end, and what was at the end now comes first. Then the mirror is used as a transmitter array, and the time-reversed signals are sent out again.

Fink's reasoning was that since the wave equation is time-reversal invariant, the wave that is sent back should repeat its trajectory in reverse and focus on the original source point. His first objective was to focus sound waves on a kidney stone to destroy it, although this was followed by other applications. Once he had built the apparatus, his first experiment was conducted in a tank of water – a perfectly homogeneous medium. In the homogeneous medium, Fink developed a theory to predict the focal spot size, in other words the volume on which he would be able to focus the acoustic energy.

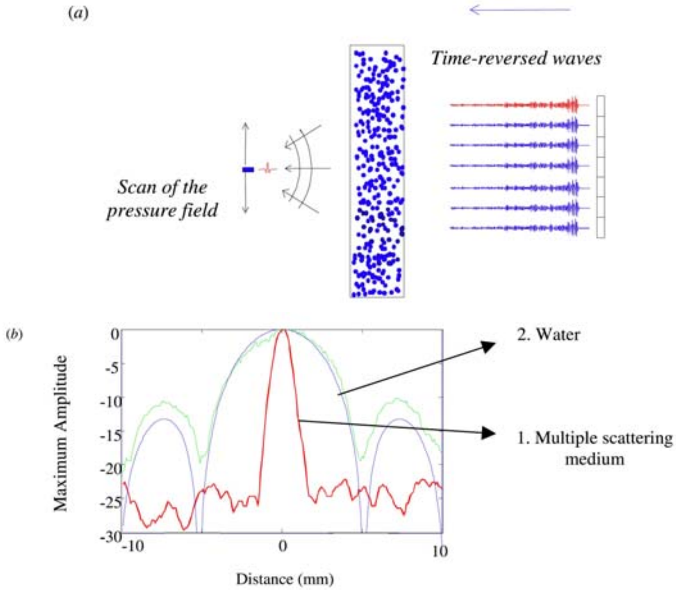


Figure 5: Time reversal – refocusing of the wave on the original source point (Tourin et al., 2000)

Conducting experiments in the water tank, he obtained focusing results consistent with the theoretical predictions for a homogeneous medium. Before attempting to focus ultrasound waves on the kidney stones of real patients, he wanted to test the technique in heterogeneous media to verify that the energy focused equally

© Cournot Centre, November 2012

well. He therefore repeated the experiment, but with the introduction of inhomogeneities into the medium, in the form of metal rods. Repeating the experiment with increasing numbers of metal rods, he observed that the focusing became more and more efficient. The more heterogeneous the medium, the better the focusing (Figure 5).

At the end of the 1990s, this experiment was presented to a CNRS (French National Center for Scientific Research) research group containing a number of mathematicians, specialists in wave propagation in random media, who were astonished to learn that the results improved as the level of heterogeneity increased. These mathematicians then carried out research to clarify the relation between the refocused signal observed during a time-reversal experiment and Green's function, that is to say, the fundamental solution to the wave equation. Green's function is what we measure at an observation point (*x*) when we emit a short pulse from a source point (*y*) in a direct propagation experiment (Figure 6). In the time-reversal experiment (Figure 7), a source at *y* emits a short pulse that propagates through the medium; the signals are recorded on the time-reversal mirror; they are then time-reversed and re-emitted, and we record the new signal obtained at *x*. Although the direct propagation and time-reversal experiments appear to be very different, we can observe a relation between the refocused signal obtained at *x* in the time-reversal experiment and the direct signal obtained at *x* from a source point at *y*.

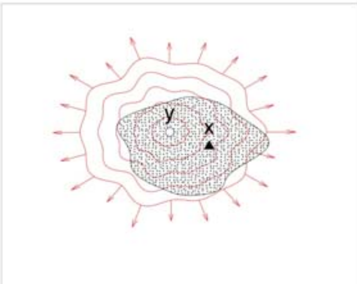


Figure 6: Green's function – signal measured at *x* when a source at *y* emits a short signal.

© Cournot Centre, November 2012

These are two experiments that do not appear to have anything in common, and yet they are equivalent – except for one symmetrization – since the signal refocused at *x* in the time-reversal experiment is a slightly distorted version of Green's function (the causal version and the anti-causal version are superimposed).

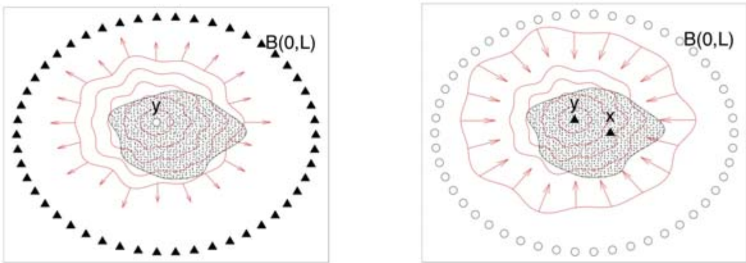


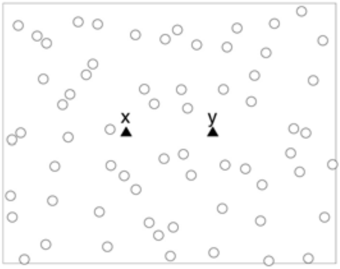
Figure 7: Time reversal – from the emission of a short signal at *y* to the refocusing of the time-reversed wave

The researchers in mathematics (Papanicolaou, myself and others) worked on these problems from the end of the 1990s up until 2005–06, seeking to explain why the experiment worked, and why it worked even better in a random medium (Fouque et al., 2007). Our mathematical analysis of time reversal for ultrasounds, that is, the study of the time-reversal operator, led us to discover that disorder does indeed increase the focusing capacity of the waves. These results were presented to the CNRS research group, which included specialists in very different fields: optical scientists, acousticians, mathematicians, and also seismologists. The latter observed that it was not possible for them to build a time-reversal mirror of the vast size needed to send seismic waves back through the Earth's crust, thereby raising the question of the utility of these ideas for their discipline.

The answer came in 2005: the time-reversal operator has the same form as a cross-correlation operator, which is used on stationary random signals (the signals used by mathematicians to model noise). This observation is of little interest in ultrasound imaging, but of great interest in seismology, where phenomenal quantities of noise are available (networks consisting of several thousand seismometers), together with years of seismic noise recordings waiting to be used.

© Cournot Centre, November 2012

The principle is the following (Figure 8): the circles represent sources of noise (○) emitting stationary random signals with zero mean and with any power spectral density – it could even be white noise. These continuously emitted signals propagate through the medium and are recorded by two sensors at points *x* and *y*. Figure 8 (right-hand side) shows the signals recorded at *x* and at *y*: we observe that noise is emitted, and noise is then recorded.



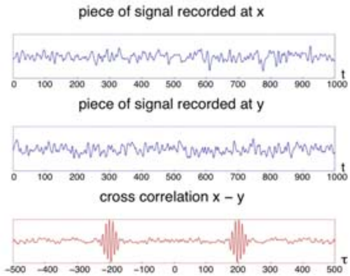


Figure 8: Cross correlation of ambient noise: configuration (left-hand figure), signals recorded and cross correlation of signals (right-hand figure)

The question is whether one can extract information from the recordings made at two observation points, and in particular whether it is possible to estimate the propagation velocity of waves between *x* and *y*. The answer to this question depends on the calculation of the cross correlation:

$$C^{T}_{x,y}(t) = \frac{1}{T}\int_{0}^{T} u_{\mathrm{x}}(s)u_{\mathrm{y}}(t+s)ds$$

We take the signal $u_x(s)$ recorded at *x*, then the signal $u_y(s)$ recorded at *y*, shift the latter by a time interval *t*, then multiply the two together and calculate the average over the whole time window of the observation. If we have a large quantity of noise, that is to say, a very long observation time, then we can obtain a relation between this cross correlation and the signal recorded at *x* in the time-reversal experiment with the source point at *y* (Figure 7). As mentioned above, there is a relation between the signal recorded in the time-reversal experiment and the signal recorded in the direct experiment with source point at *y* and observation point at *x*.

© Cournot Centre, November 2012

Consequently, there is a relation between the Green's function recorded in the direct propagation experiment and the cross correlation of the noisy signals recorded in the experiment with sources of stationary noise. As we know that the Green's function has a wave-front – a main peak – which gives the travel time, we can use this relation to propose a method of estimating the travel time between *x* and *y* based on the cross correlation of noisy signals. Finally, if we know the travel time between *x* and *y*, then to the extent that we also know the distance between *x* and *y*, we can obtain an estimation of the wave velocity between *x* and *y* by dividing the one by the other.

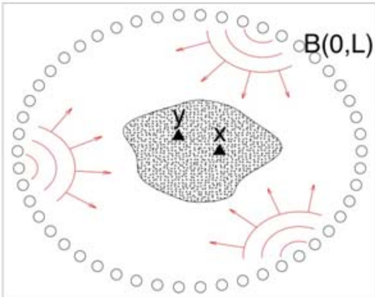


Figure 9: Cross correlation of ambient noise signals recorded at *x* and *y*

In the cross correlation experiment (Figure 9), there is no time-reversal mirror, but we do have sources of noise (the circles) emitting waves that are propagated and then recorded at *x* and *y*. We then compute the cross correlation of the signals recorded and obtain a relation with the Green's function between *x* and *y*, except for one symmetrization and one convolution, because the power spectral density of the sources remains. If the sources emit white noise, the equality is perfect; if they emit "coloured" noise, that is to say, the power spectral density is not flat, the Green's function will be filtered by the spectral density of the sources. In other words, if there is a gap in the noise spectrum within a certain range of frequencies, then it will not be possible to estimate the Green's function within that range of frequencies.

The implementation of these ideas took time, because it was necessary to understand the properties of a mathematical object that constantly reappeared: the time-reversal operator. This operator models the time-reversal experiment and can be defined as the result of the convolution of two Green's functions, one of which is

© Cournot Centre, November 2012

time reversed. This operator also comes into play in the modelling of cross correlations of noisy signals. This is because a correlation is similar to the composition of a convolution with a time reversal. The implementation of this technique of estimating travel times by computing cross correlations works remarkably well in the field of seismology.

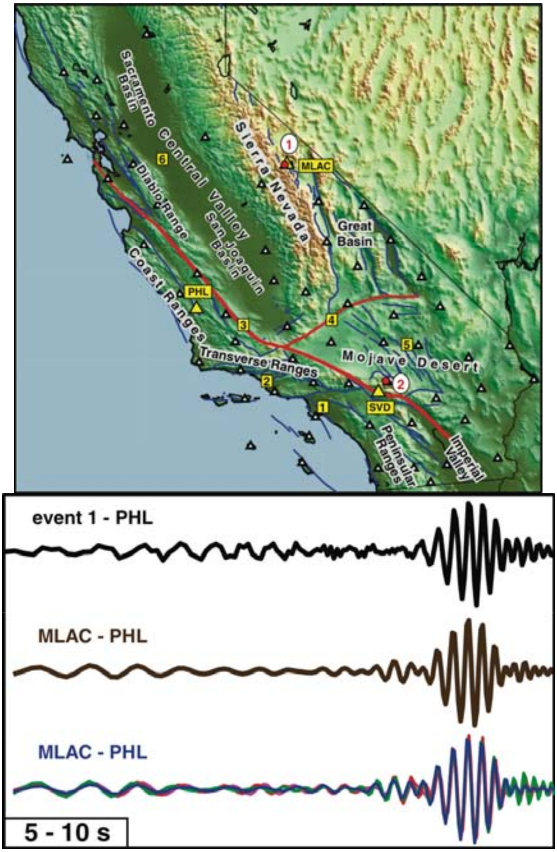


Figure 10: Comparison between Green's function and cross correlation between two sensors (Shapiro et al., 2005)

Figure 10 presents the first results obtained for Southern California (Shapiro et al., 2005). This figure compares three plots that resemble each other, and yet they

© Cournot Centre, November 2012

were not obtained in the same way. The first plot is the signal recorded by the PHL seismometer of an earthquake centred on point 1. The second plot is the cross correlation of the seismic noise recorded over one year by the seismometers MLAC and PHL. In this case, point 1, which is the epicentre of the earthquake, corresponds almost exactly to the location of the seismometer MLAC. This enabled researchers to make a comparison between the direct signal produced by the seismic source and the cross correlation between the seismic noise recorded by the two seismometers MLAC and PHL. There is a perfect match. The third plot shows three cross correlations obtained with three sets of four-months' worth of signals, and we can see that the cross correlations are practically identical. The slight discrepancy shows that we need a large quantity of noise for the cross correlation to produce the same result as we get with a real, impulsive seismic source. This demonstration is what everybody was waiting for: now it was no longer necessary to wait for earthquakes to do seismic imaging; one only had to continuously record seismic noise, to consider the signals for each pair of seismometers, and then to calculate all the cross correlations.[6]

Figure 11 shows an array of about 60 sensors. The straight lines represent all the estimations of travel times that have been obtained between all the pairs of seismometers. This can then be used to draw up velocity maps of seismic wave propagation (Figure 12) that the seismologists find not only satisfactory, but actually better than those based on recordings of signals produced by earthquakes, because the latter are often high-frequency signals that do not penetrate very deeply into the Earth's crust.

Before speaking of other applications, I should specify where the seismic noise comes from. Not only in Southern California, but over the whole surface of the globe, the low-frequency components (periods of 1 to 10 seconds) come from the ocean, and more precisely from the interaction of waves with the ocean floor. This is not surprising in California, being close to the ocean, but the experiment has been carried out all over the planet – in the Alps, in Tibet – and even in Tibet the seismic noise comes from the oceans. One surprising manifestation of this origin of seismic noise was observed in California in 2005. It had been observed that most of the noise always came from the west, except for one short period when the direction of the flow

[6] In fact, seismic background noise is comprised above all of low-frequency components, whereas signals from earthquakes contain primarily high-frequency components, so the two techniques are also complementary.

© Cournot Centre, November 2012

was reversed, and that was when the hurricane Katrina hit Louisiana. Likewise, there is a seasonal effect between winter and summer in Tibet, with periods when the noise comes mainly from the Indian Ocean and others when it comes mainly from the Pacific. Now, a great many researchers work on seismic noise. A new world has opened up, where stocks of data await eager researchers who can now use them to do imaging.[7]

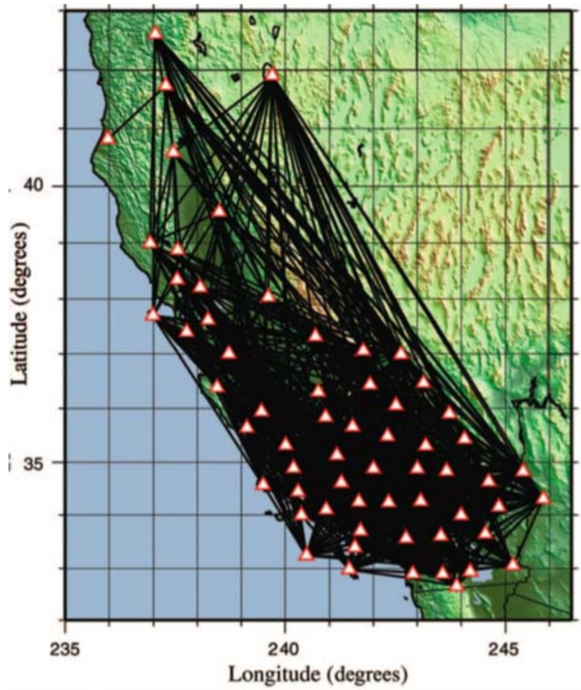


Figure 11: Estimated travel times between pairs of sensors (Shapiro et al., 2005)

Today, this research is conducted by seismologists like Michel Campillo of the ISTerre (Institute of Earth Sciences) in Grenoble. The results obtained for the imaging of volcanoes are particularly interesting. This is imaging at a finer scale, using seismic noise at higher frequencies. The advantage is speed: maps can be drawn up in a day!

[7] See the special issue of *CRAS Geoscience* (2011), 343 (8–9).

© Cournot Centre, November 2012

This was first implemented to map the Piton de la Fournaise: by carrying out daily imaging, starting with data from previous years, the seismologists realized that they were capable of such precise imaging that they could see fluctuations in wave velocity of less than one in a thousand, forewarning of an eruption. They were thus able to predict the most recent eruption of the Piton de la Fournaise three weeks in advance (compared with a previous maximum of 48 hours). So the method works at different scales. It could almost certainly be useful in the search and monitoring of gas or oil deposits or CO_2 storage.

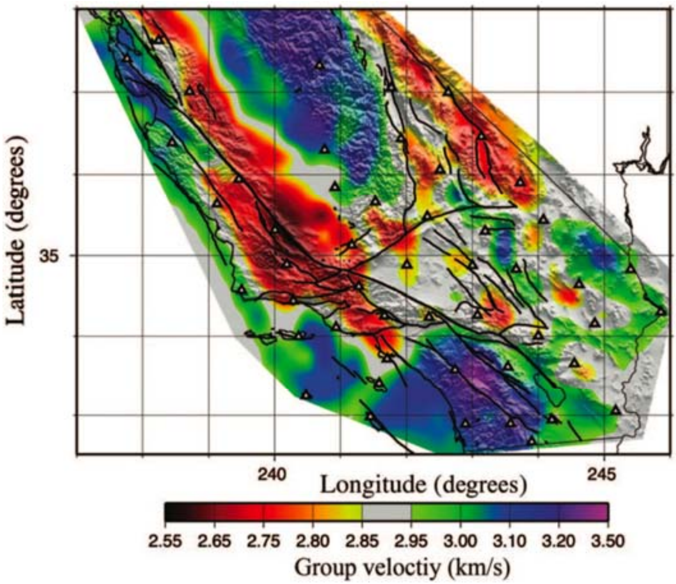


Figure 12: Estimated sound speed based on estimated travel times (Shapiro et al., 2005)

This research has also been pursued on a more theoretical level by mathematicians. Figure 13 shows a very recent development in imaging, where instead of estimating the sound speed, the aim is to locate reflectors – more or less distant objects.

The first diagram in Figure 13 shows noise sources, five sensors and a reflector buried in the medium (diamond in Figure 13). The question is whether one

© Cournot Centre, November 2012

can find the reflector by using the signals emitted by the noise sources and recorded by the five sensors. The answer is yes, thanks to the method described above. By calculating the cross correlations, pair by pair, of all the pairs of sensors, we obtain a matrix of cross correlations. The image obtained in Figure 13 (bottom) results from the migration of the 10 available cross correlations. With the signals emitted by the ambient noise sources, we can obtain an image and determine the location of the reflector. Papanicolaou and I have been working on the imaging of reflectors for about two years now. It is not as straightforward as the estimation of travel times, because we have to use migration techniques that are rather complicated to apply to these sorts of matrices of results. So this is still a work in progress.

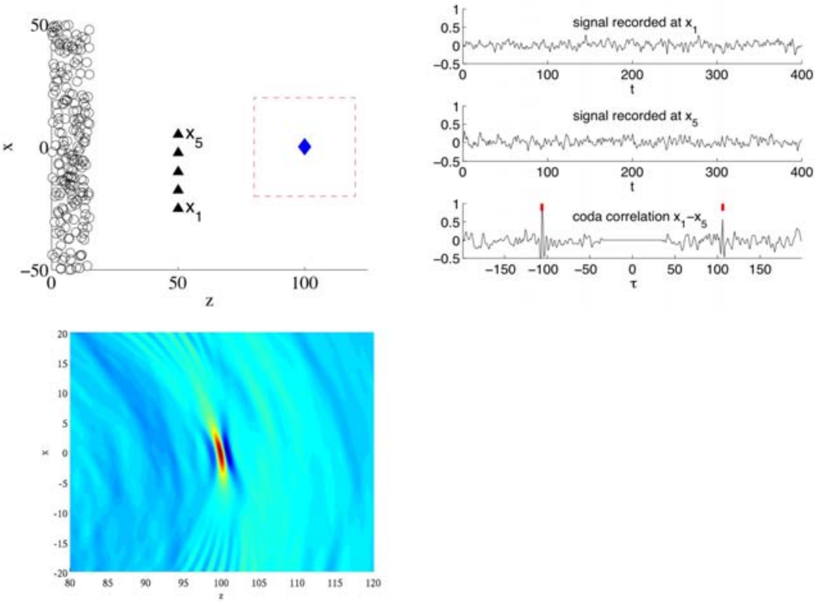


Figure 13: Ambient noise imaging of a reflector

Other researchers are working on various extensions of the method, including applications to unexpected domains like lunar seismology. In fact, NASA has stored almost one years' worth of signals recorded by an array of four seismometers installed on the moon's surface during the Apollo 17 mission. By computing cross correlations, it has been possible to extract information from these old recordings

© Cournot Centre, November 2012

(Larose et al., 2005). The seismic noise on the surface of the moon derives from the large thermal contrasts that crack the rocks, constantly producing noise on the surface. It is also possible to extend the technique to imaging the internal structure of the sun, a study known as helioseismology. In fact, the experiment turns out to be easier to do on the sun than it is on the Earth. Because of its solidity, the waves that propagate through the Earth are not acoustic but elastic, entailing different modes and speeds of propagation. Since the sun is fluid, all the waves are acoustic, and there is only one propagation speed. By observing the pulsations of the sun, through the Doppler effect, we now know more about the internal structure of the sun than about that of the Earth. We obtain many more details, and in three dimensions, moreover, rather than a simplified image with spherical symmetry. Some researchers are applying this technique to other stars, practising what they call "asteroseismology", and thus producing images of the internal structure of distant stars.

Papanicolaou and I are now working on a new domain of wave imaging, studying the propagation of microwaves. These are no longer acoustic waves. We are seeking to do passive imaging in the range of frequencies around 2 gigahertz, because this range contains a lot of noise from Wi-Fi terminals, mobile phone stations and mobile phones, creating a sea of microwaves in which we are constantly immersed. We are trying to determine, with some results already, whether it is possible, for example, to perform geolocation and then imaging based on this ambient microwave noise. The use of noise in imaging has enabled a major breakthrough, the future developments of which will continue to surprise us.

© Cournot Centre, November 2012

References

Aki, K., and B. Chouet (1975), "Origin of coda waves", *J. Geophys. Res.*, Vol. 81, pp. 3322–42.

Fouque, J.-P., J. Garnier, G. Papanicolaou and K. Solna (2007), *Wave Propagation and Time Reversal in Randomly Layered Media*, New York: Springer.

Garnier, J. and G. Papanicolaou (2009), "Passive sensor imaging using cross correlations of noisy signals in a scattering medium", *SIAM J. Imaging Sciences*, Vol. 2, pp. 396–437.

Larose, E., A. Khan, Y. Nakamura and M. Campillo (2005), "Lunar subsurface investigated from correlation of seismic noise", *Geophys. Res. Lett.*, Vol. 32, L16201.

Shapiro, N., M. Campillo, L. Stehly and M. Ritzwoller (2005), "High-resolution surface-wave tomography from ambient seismic noise", *Science*, Vol. 307, pp. 1615–18.

Tourin, A., M. Fink and A. Derode (2000), "Multiple scattering of sound", *Waves Random Media*, Vol.10, pp. R31–R60.

© Cournot Centre, November 2012