

Seeing through the donut holes

a conversation between a young and an elderly seismologist

(Karin Sigloch & Guust Nolet)

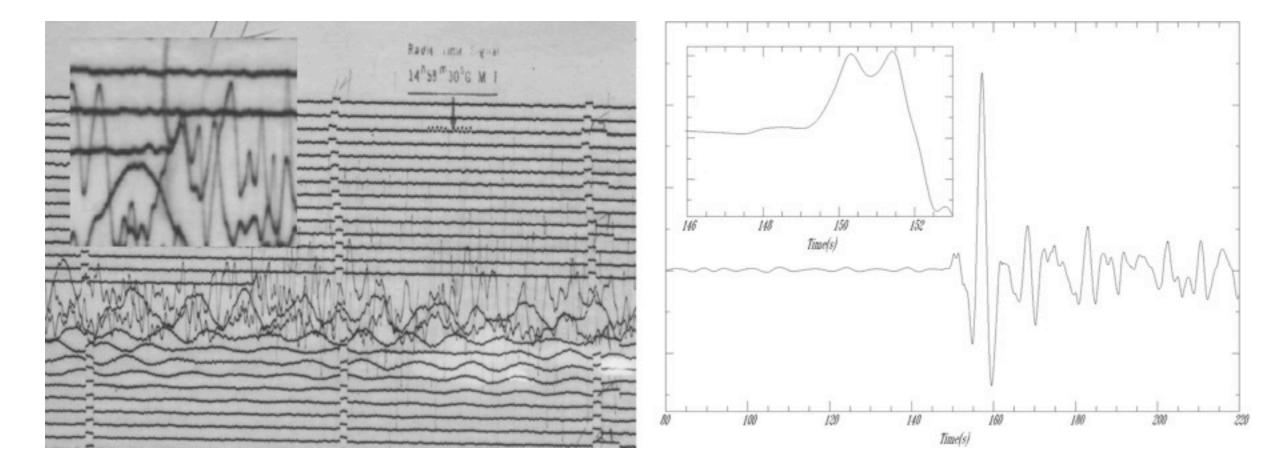
Monday, 21 May 2012

G: Hi Karen, long time no see, remember me?

K: I think so, but that was ages ago. Have you retired?

G: Almost, though I still look at seismograms. I don't keep up with this new stuff, though. What's all that hoopla about donut holes?

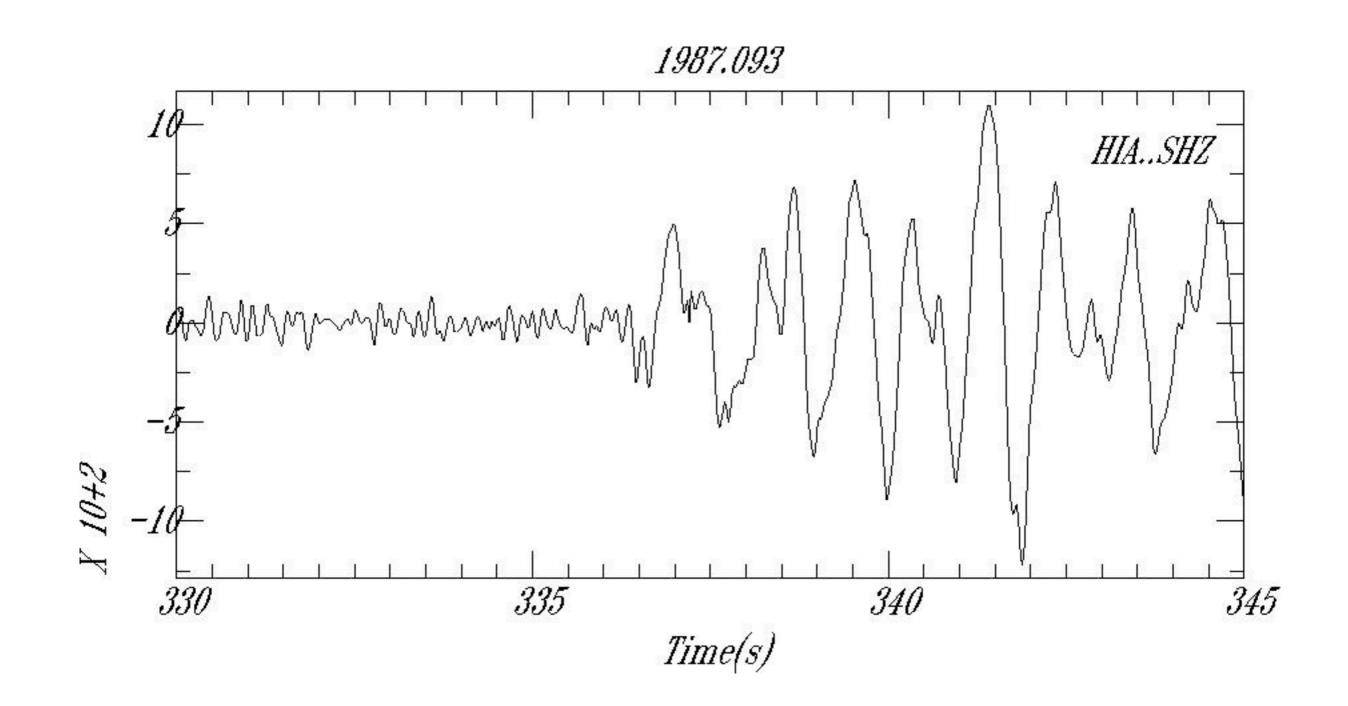
onset time



Monday, 21 May 2012

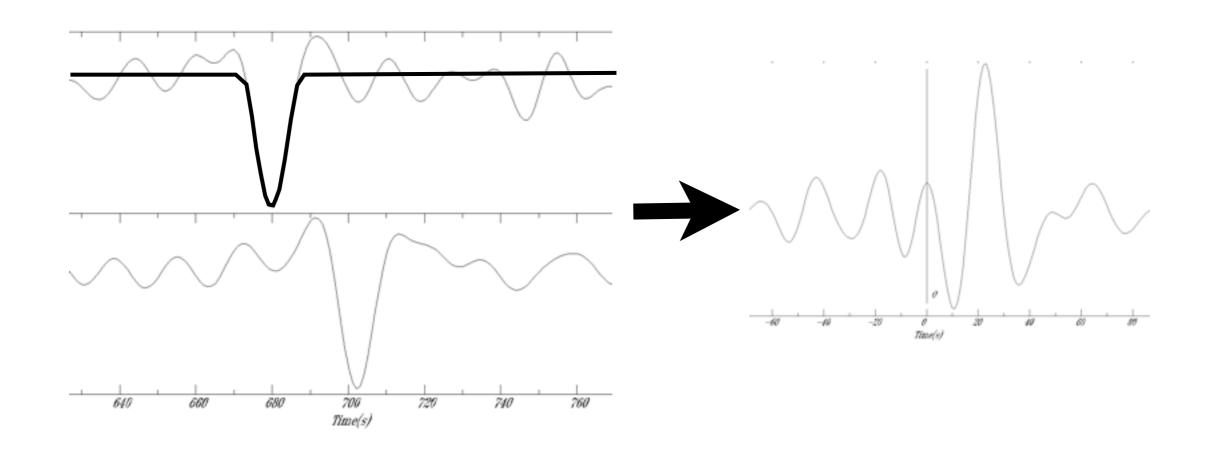
K: I guess you're still picking onset times? That is so out of fashion.... ever since we have digital seismograms.

G: Mmm, yes. But what's wrong with it? I now pick onsets with SAC on a screen, so it is actually very easy.



K: Yes, as long as there is no noise, and you can actually see the very first arrival. But even then, what have you got? Just one piece of information. You forget that there may be thousands of samples in your SAC file. What do you do with them? Throw them away? What a waste! 1024 or more samples and only one bit of information....

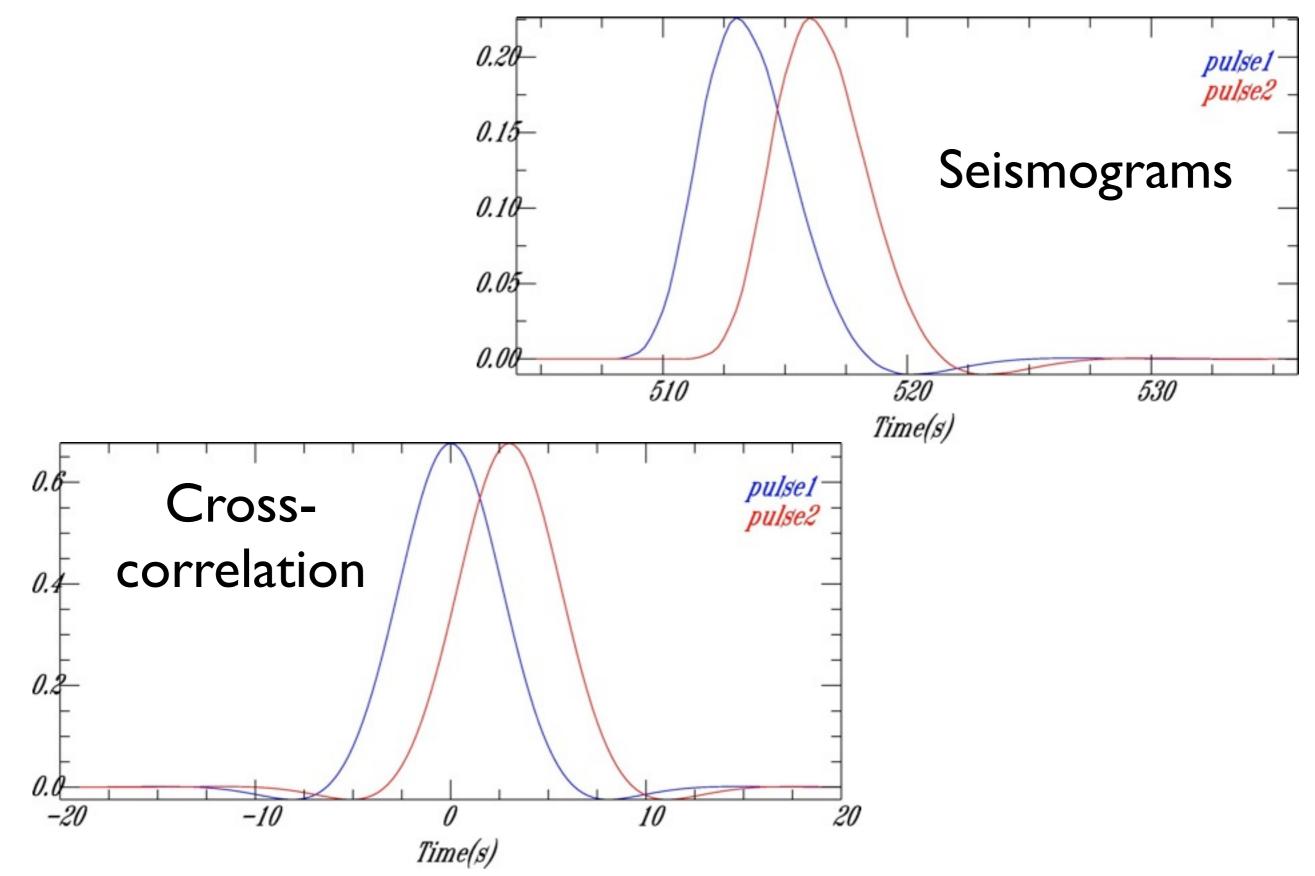
cross-correlation



$$C_{uv}(t) = \int u(\tau)v(\tau - t) \, \mathrm{d}\tau$$

Monday, 21 May 2012

G: No, no. I can pick later arrivals. Especially with cross-correlation that has become quite easy, even in the presence of noise. It gives you a very accurate time measurement. K: Yes – in fact I do that myself. But that implies that you cannot use ray theory!

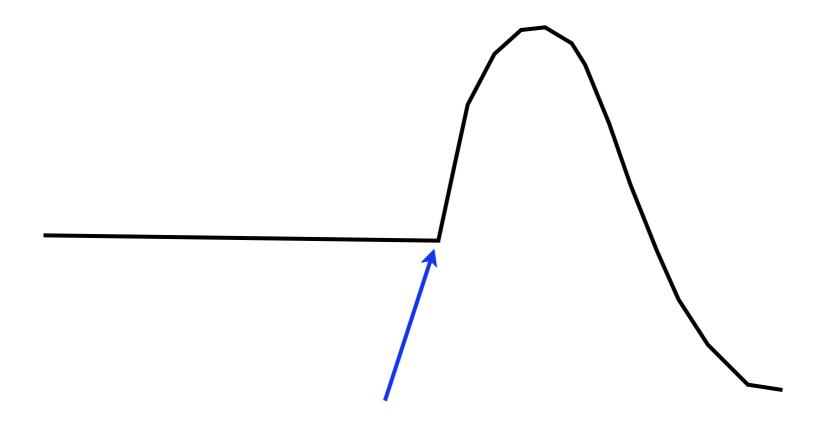


G: I'll show you! Here I have a blue fast pulse and a red one that has been delayed by 3 seconds. The onset times thus differ by 3 seconds...I know because I synthesized them myself. And when I take the cross-correlation with the blue one I get precisely 3 seconds for the red one. Since I measure the same delay, I can evidently use the same theory.K: But how did you compute them?G: I low-passed a blue pulse, then shifted it 3 seconds to get the red one. You don't have to be a rocket scientist to do that.

K: So there was no physics in your calculation, just math?

G: What do you mean?

The trouble with onsets



- all frequencies arrive at same time (zero phase)
- no frequencies have been attenuated away
- (and we are not even talking about instrument response...)

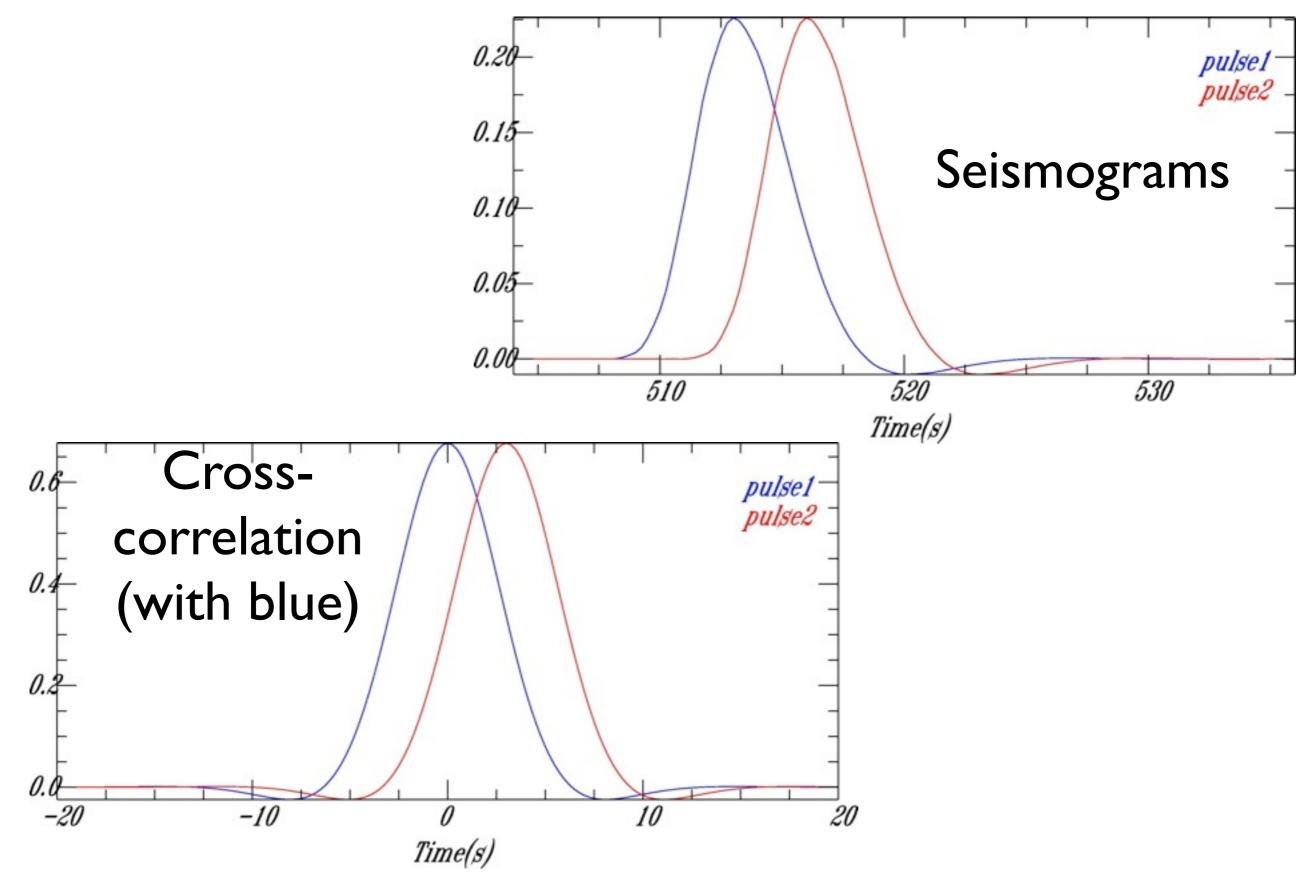
have attenuated (click) and I am not even talking about the instrument....

Monday, 21 May 2012

K: well, let us first see what physics we need to get a sharp onset, because ray theory assumes you can actually detect the minimum arrival time. Remember that you can derive the eikonal equation by minimizing the travel time?

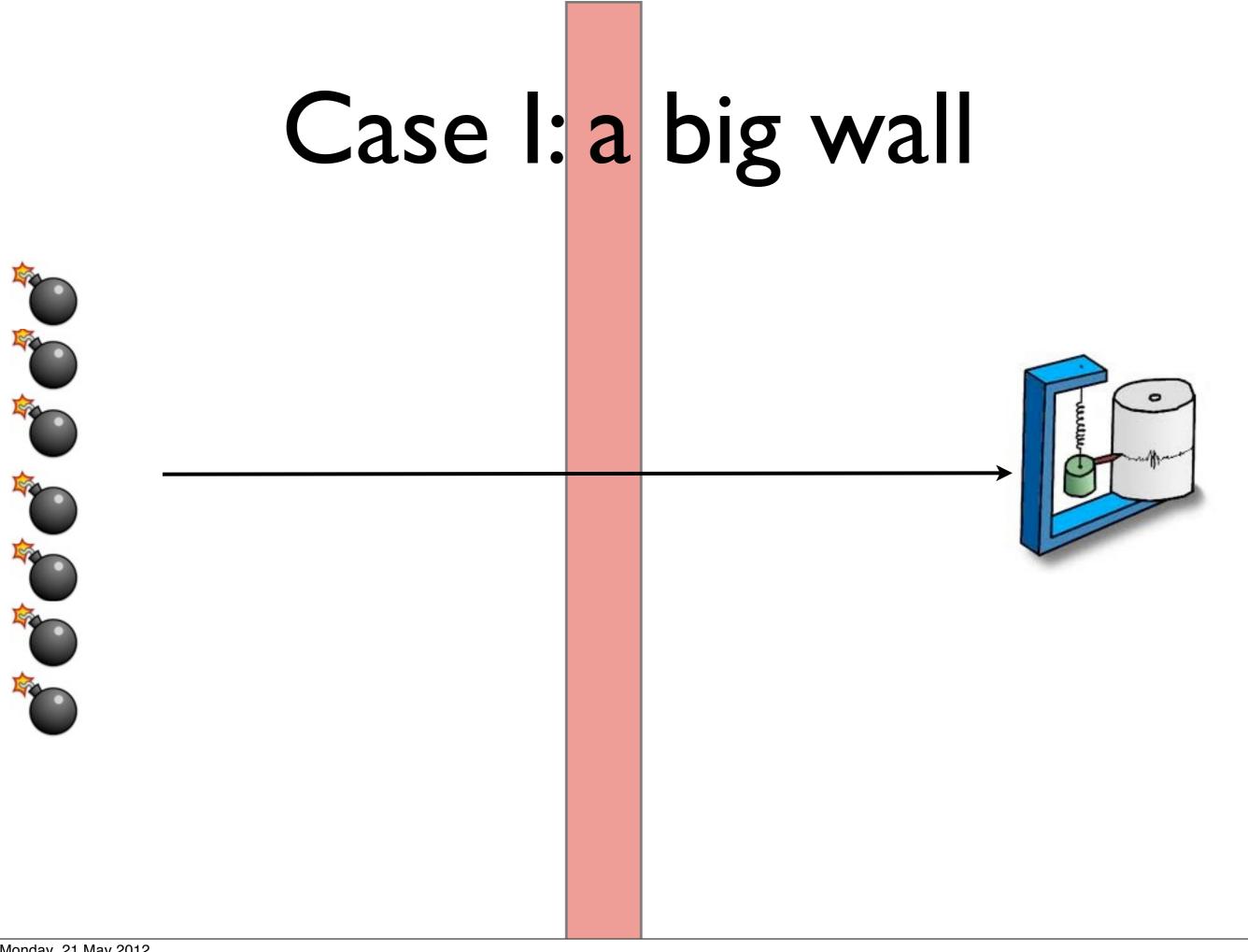
G: of course – I taught you so!

K: well (click) you want all frequencies present in the onset – so they must all be zero phase (click) and should not



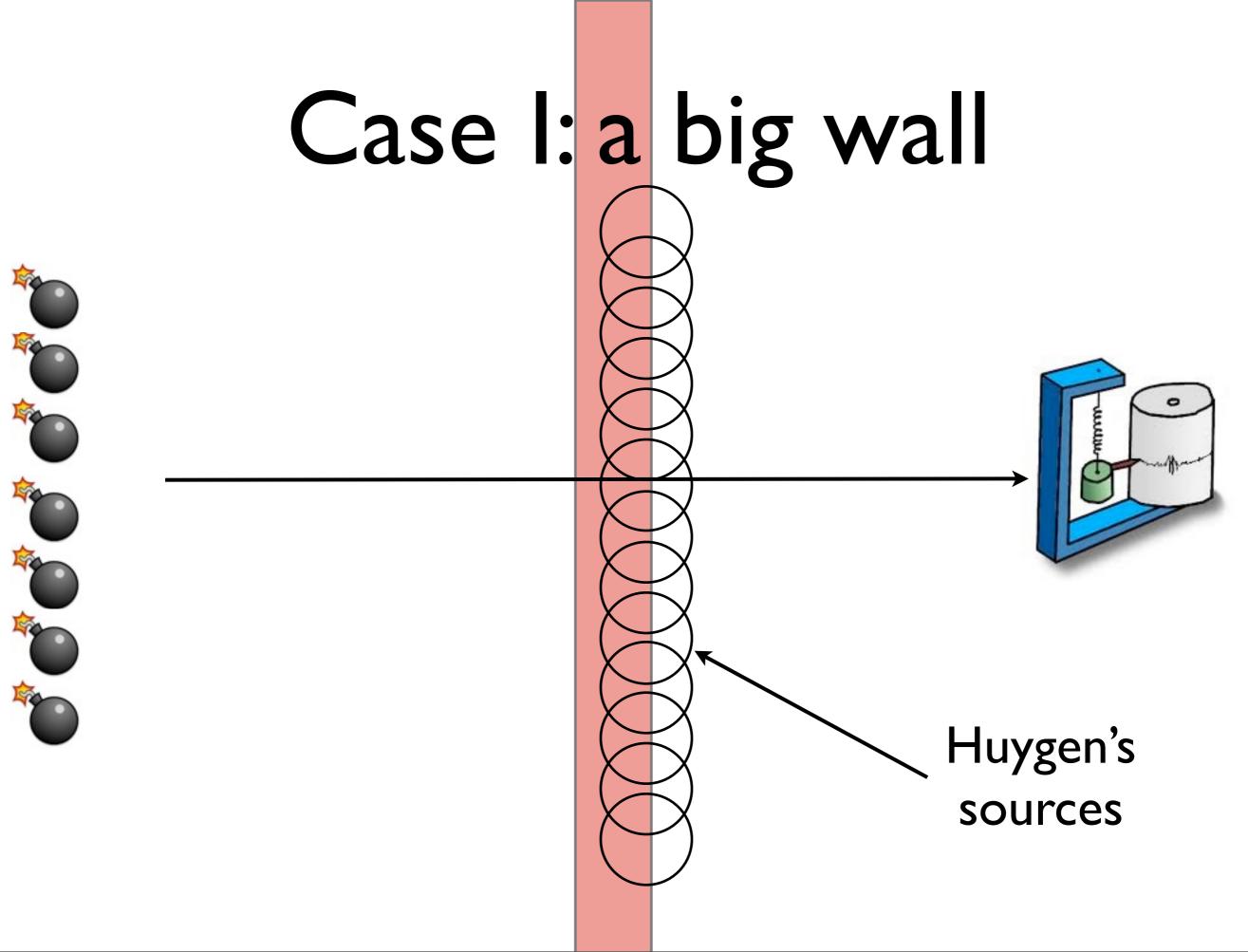
G: my point is that I can still measure the 3 seconds delay even if the onset is not so sharp, by using cross-correlation:

K: but some heterogeneity must be causing your delay, and cause an effect that is not as predicted by ray theory! Let us look at some very simple cases.



K: Take the case of a plane wave hitting a big wall. There is no way the wave can go around it. There is no diffraction of any kind, just a slowdown for the wave, no matter what path it takes to go to the seismometer, it has to go through the wall.

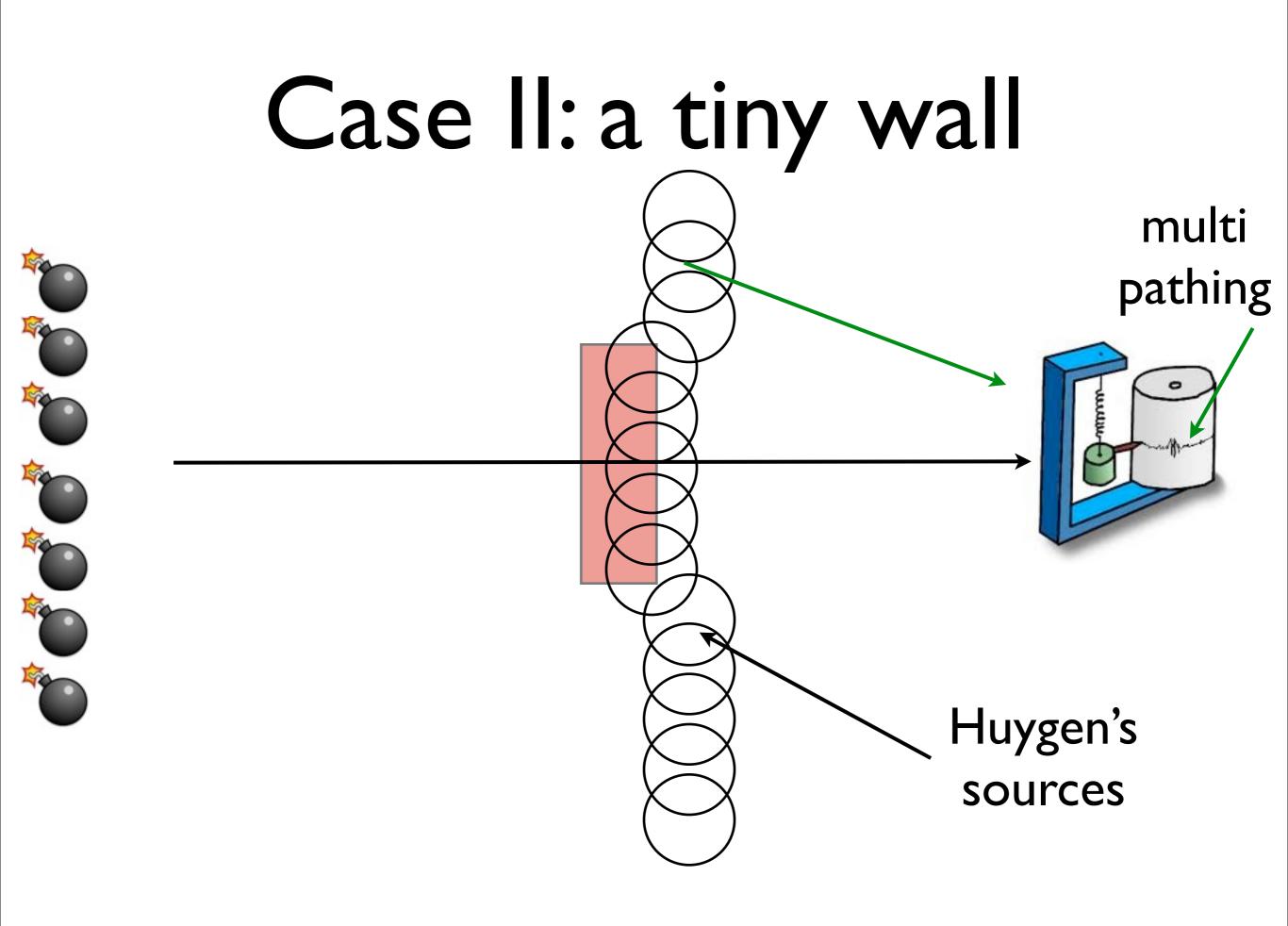
G: Mmm, yes. I can see that. You mean to say that, if I apply Huygen's law and place little imaginary sources on the wavefront when it comes out of the wall, all these sources are delayed by the wall.



K: Yes, these Huygen's sources would haven been closer to the seismometer if it was not for the wall. But the seismometer does not see a difference apart from that delay. G: Actually, in the paper by Dahlen et al they prove that a cross-correlation time predicted with this banana donut stuff converges to the ray theoretical solution for the delay

time.

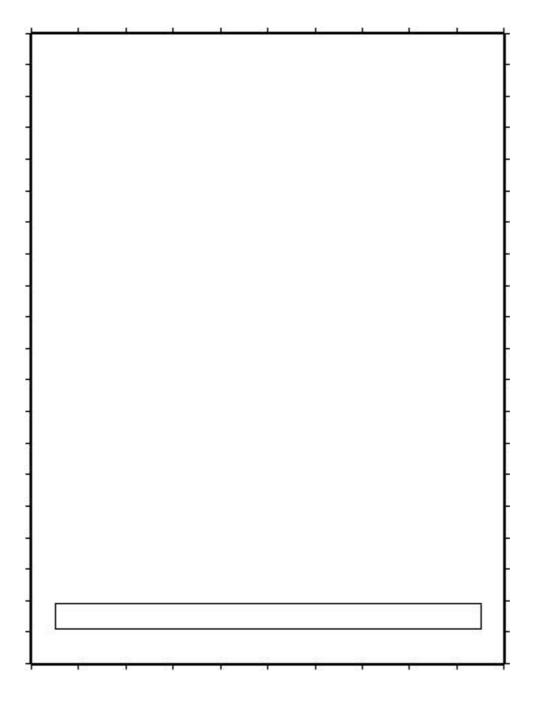
K: precisely – the waveform stays the same, the delay measured by cross-correlation is the one predicted by ray theory. But now look what happens if the wall does not extend to infinity.



K: At this point some of the Huygen's sources continue at their usual speed. They also create wave energy in the seismometer location. The detour they make is small, certainly if the instrument is far away, so they are practically not delayed.

G: but then you have a mix of waves that arrives in the seismometer. We have multipathing.... K: Yes. In this case the diffracted wave and the direct wave add up to the signal you see.

Plane wave



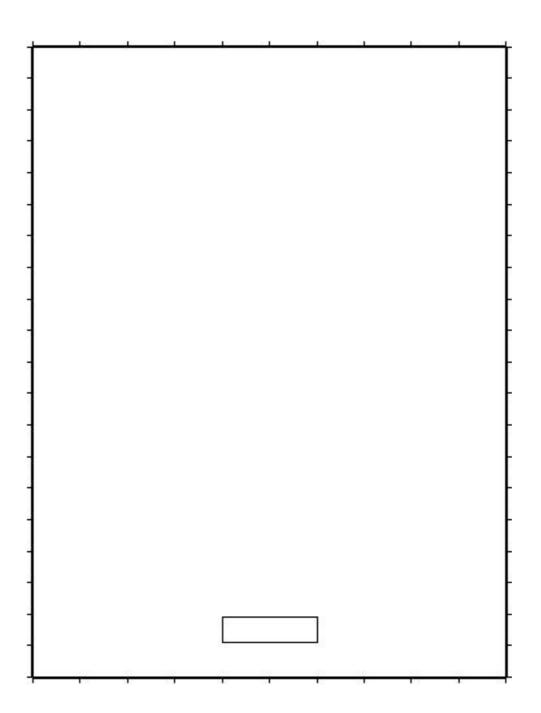
Monday, 21 May 2012

K: Here is a little movie to show the same thing. First a big wall. Notice how the wavefront remains flat.

G: But even here I see some multipathing, there are waves coming in at later time, apparently from the sides of the wall!

K: Yes, that is important if we filter: for low frequencies the window will be large enough so that these arrivals also influence the cross-correlation delay.

Plane wave



Monday, 21 May 2012

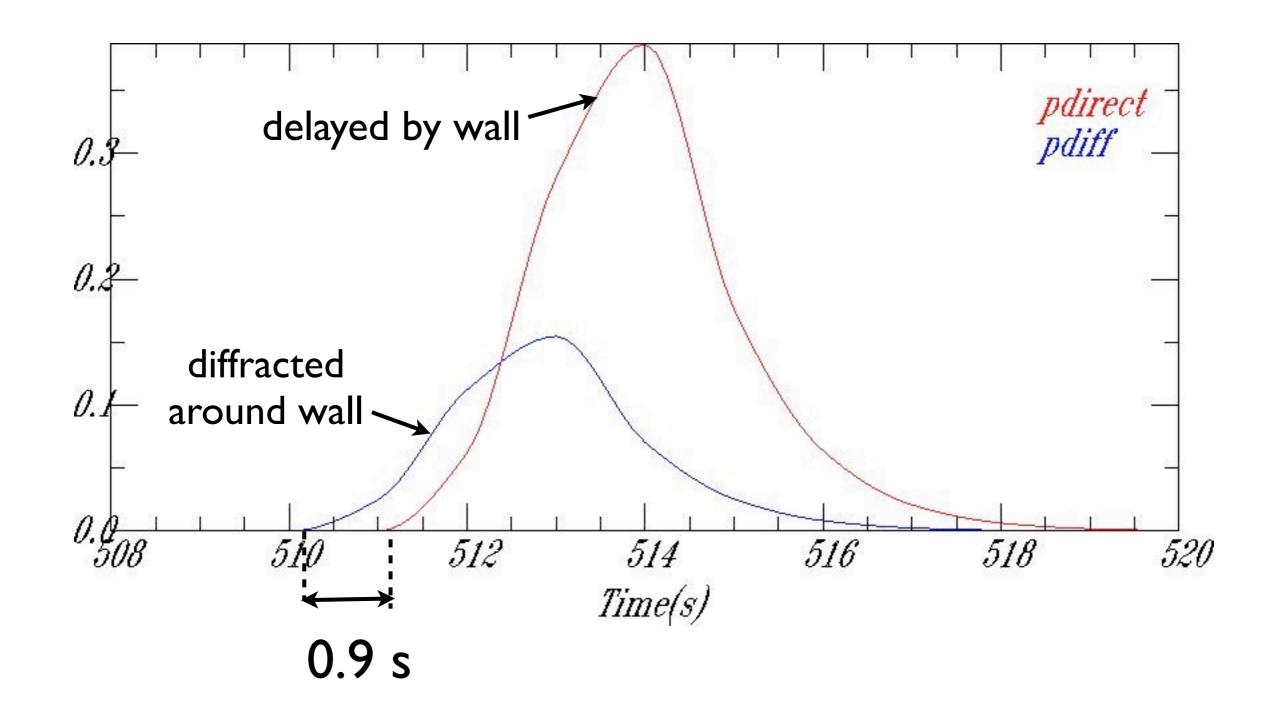
K: But if the box at the bottom is small wall it only

slows the wavefront down locally. Right after passage, one can see the delay. But as the wave progresses, the delay disappears.

G: I can see the diffracted wave filling in the delay! Can you do the movie again and stop just before it hits the other side?

- K: sure (stop movie just before the end)
- G: It looks as if the effects are much more dramatic near the side than at the center?
- K: Yes that is an effect of the donut hole: cross-correlation delays are minimal along the raypath, and larger as one moves away from it. But let us look at this from a waveform point of view.

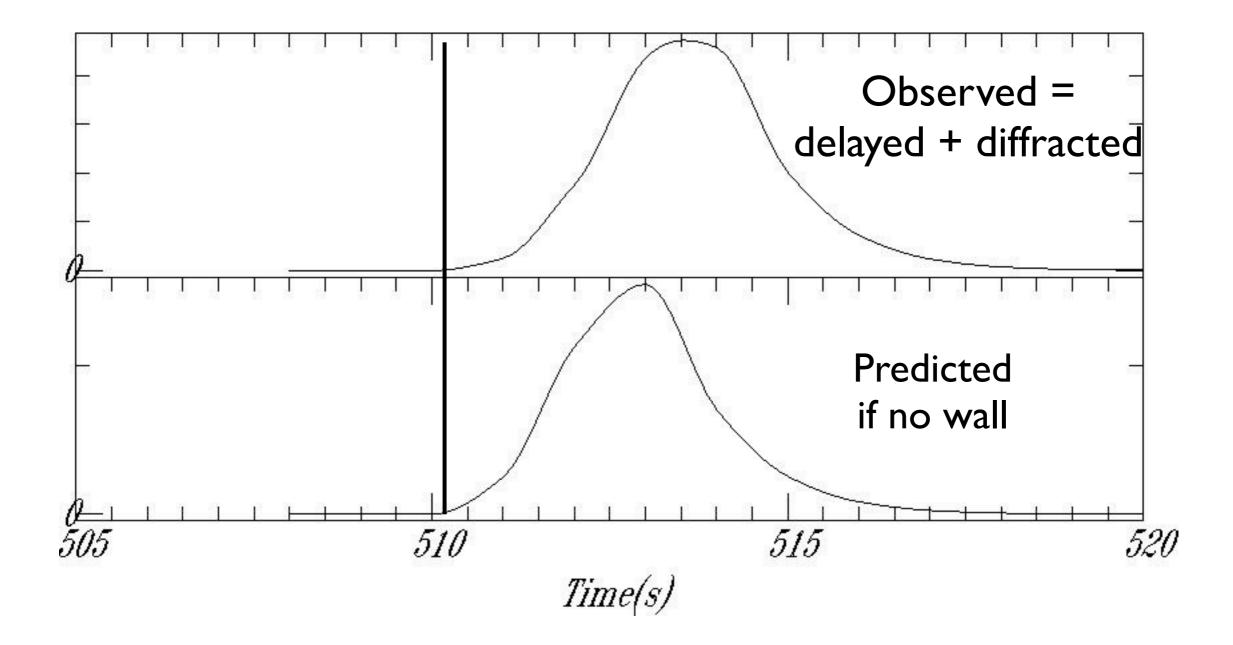
Multipathing around small wall



Monday, 21 May 2012

K: here we take the case that the diffracted wave (the blue one) has a much smaller amplitude than the direct wave – the red one. The red wave is delayed by 9/10 of a second. G: you show them separately, but on the seismogram they are added, isn't it?

X-correlation with different waveshapes

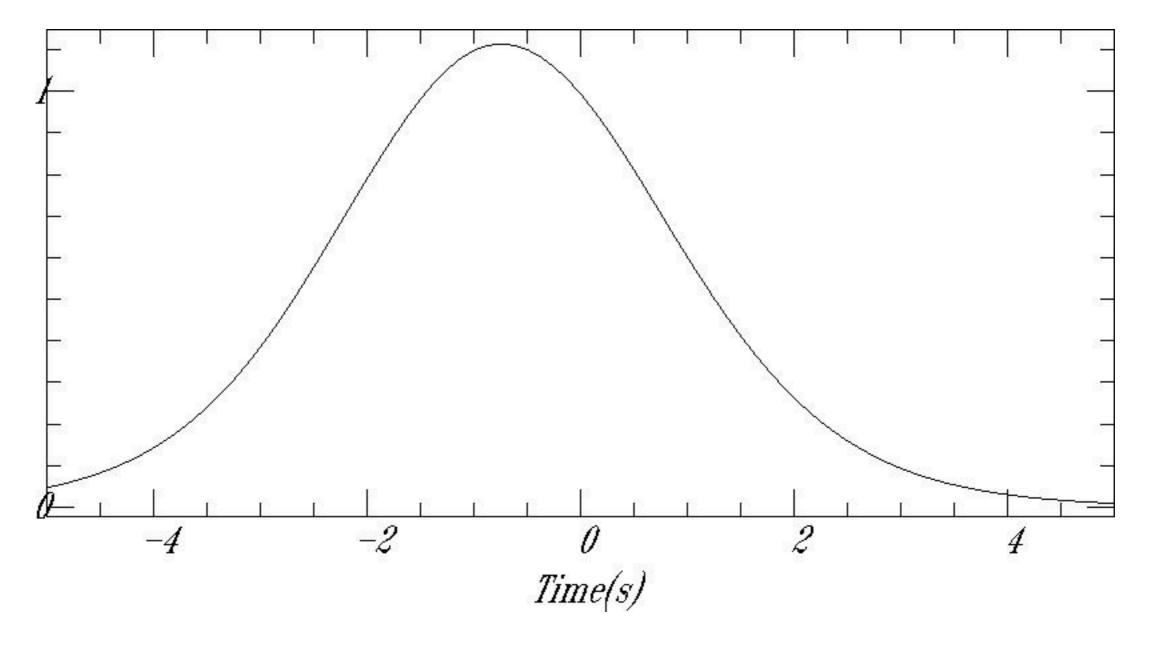


Monday, 21 May 2012

K: Yes. Here you see at the top the actual signal that we observe and that has the diffracted as well as the direct wave in it. At the bottom is a synthetic seismogram predicted for an Earth with no wall to delay the wave.

Now, I drew a line through the onsets. These are the same if we may assume that the wall is small enough with respect to the ray length so that the diffracted wave is not noticeably delayed. G: Mmm. I see. But the maxima don't come in at the same time. And the waveshape is different.... But that means I cannot do a cross-correlation!

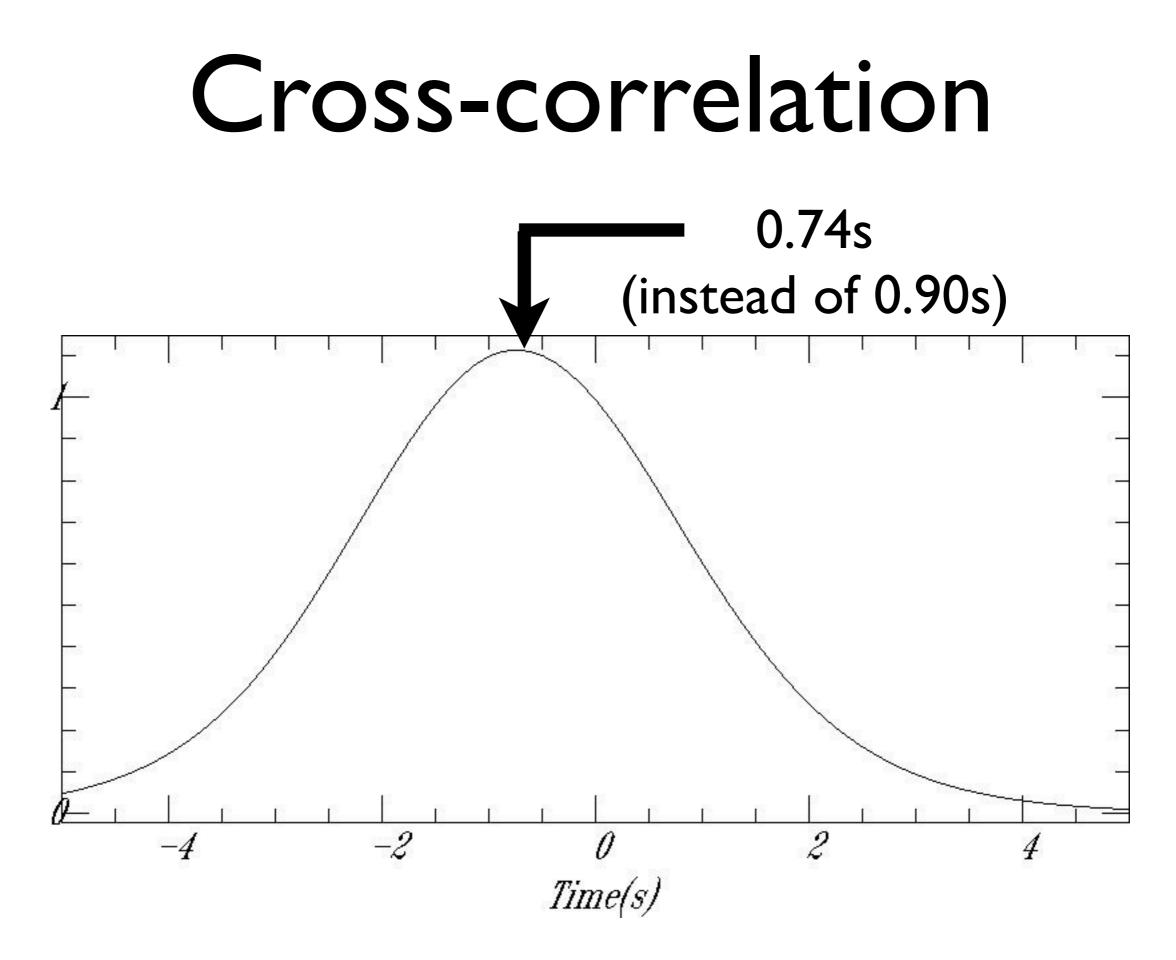
X-correlation with different waveshapes (2)



Monday, 21 May 2012

K: Well your SAC program is not going to forbid you to cross-correlate two signals... So let's do that. Do you remember the delay of the direct wave?

G: My memory is getting weaker, but numbers I still recall! 9/10 of a second was the delay that the wall gave the direct wave.



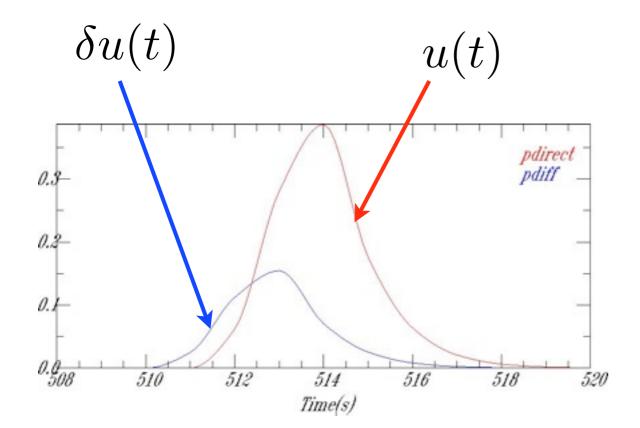
K: Yes. And the cross-correlation gives a delay of only 0.74, so we lost about 20% of the delay in this measurement. This loss of signal is called wavefront healing.

G: Shoot - that means the cross-correlation makes big errors and is useless...

K: Not necessarily. You were only expecting 0.9 because ray theory told you so. What if you move on to some better theory?

G: Isn't ray theory the best there is for a body wave?

Perturbing the x-correlation



 $\delta T_{\text{x-cor}} = -\frac{\int \dot{u}(t)\delta u(t)dt}{\int \ddot{u}(t)u(t)dt}$

Monday, 21 May 2012

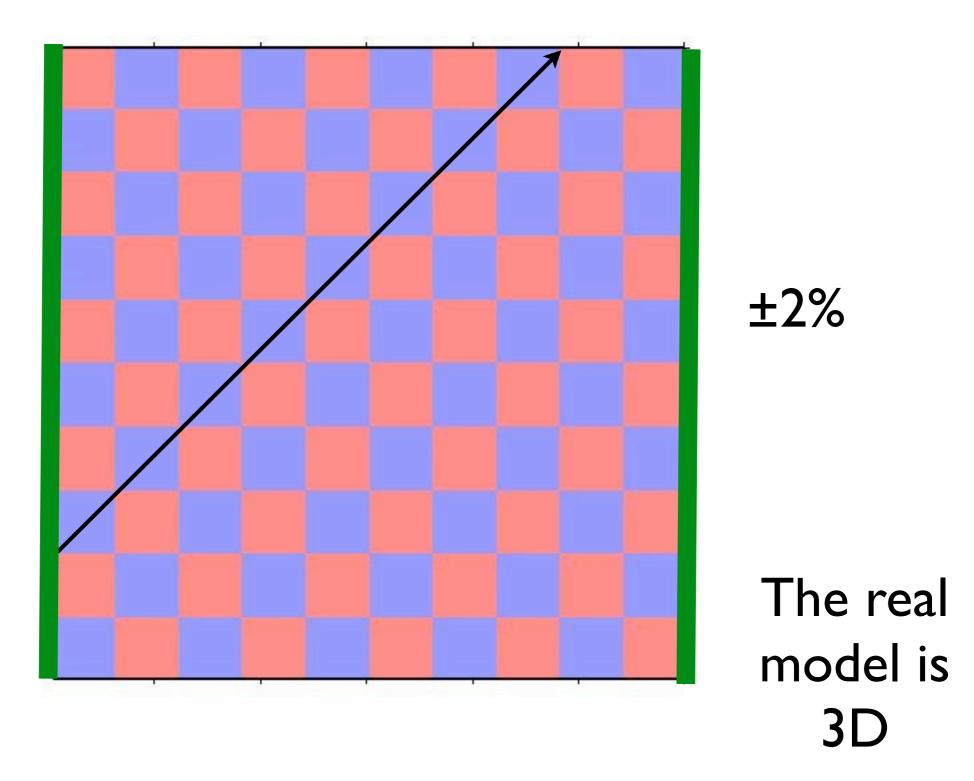
K: No, it does not predict the change in wave shape. Suppose we have an unperturbed wavefield u(t) [click] and we add to this a small perturbation delta u(t) [click]. We can then find out how the cross-correlation delay changes to first order [click].

G: Mmm... interesting... So it is linear with delta u. Of course you cannot go too far with this,

the linearity must break down at some point.

K: Yes of course; but the early numerical experiments by Shu Huei Hung showed that for a small spherical anomaly one could easily go to 6% anomalies and higher. Recently Diego Mercerat looked at a 3D checkerboard.

Borehole-to-borehole test

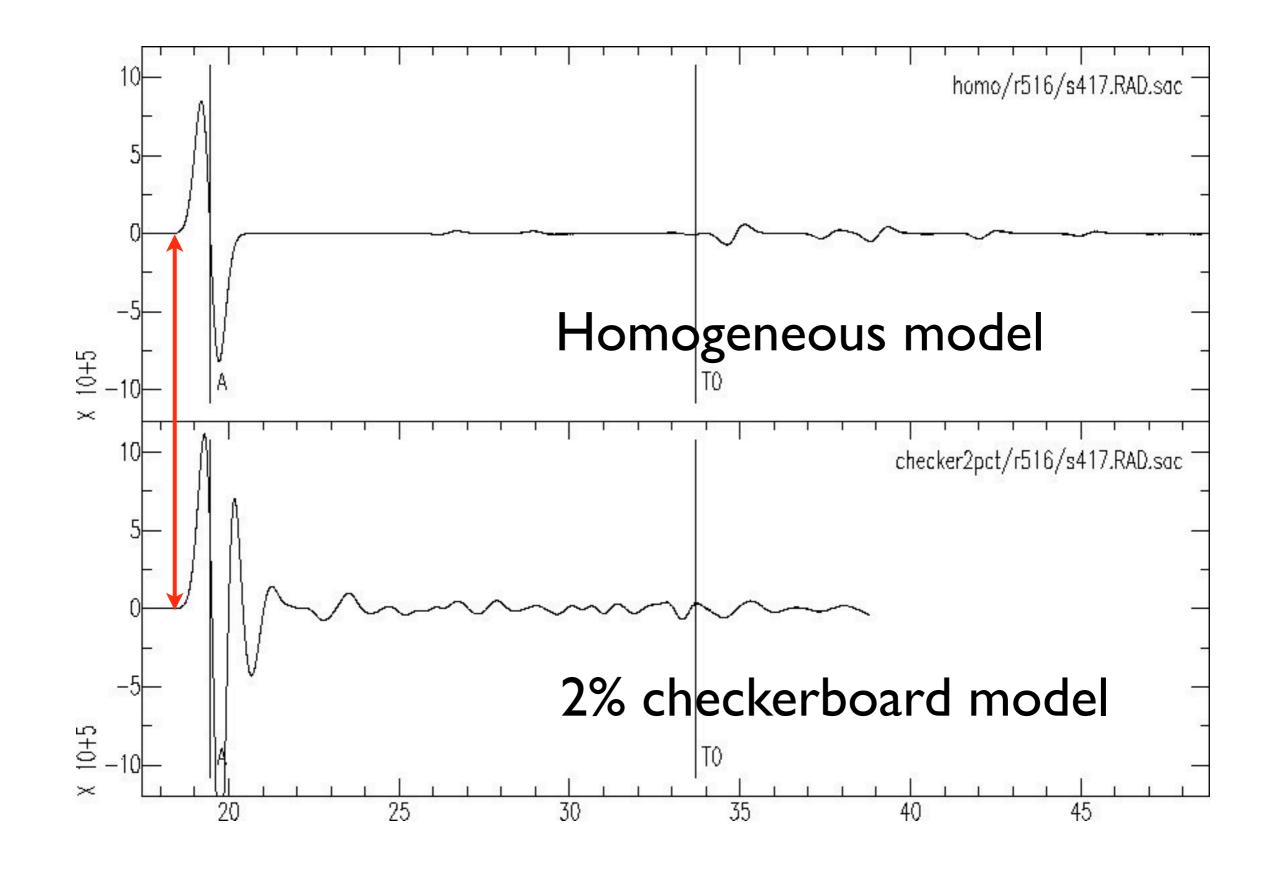


Monday, 21 May 2012

K: this 3D checkerboard has +/-2% variations in velocity. Some rays - like the one I show here - see only fast fields, others only slow, most see a mix. Diego, however, computed seismograms, not times. He got the times by cross-correlating with synthetics in a homogeneous model.

G: So you would expect travel times to be between +/-2% if ray theory is valid. What about reverberations between all those anomalies?

K: To see how they influence the linearity of cross-correlation delays is part of the test. Let us look at the seismogram along this fast raypath.

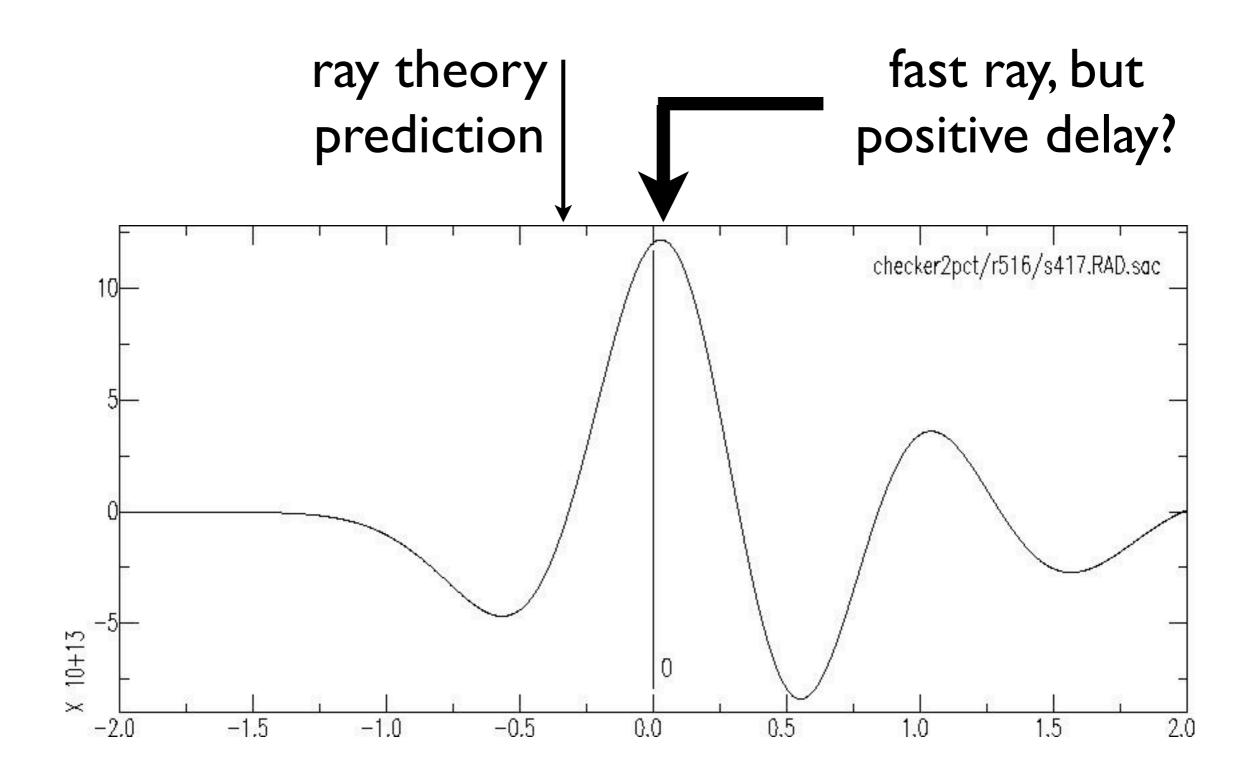


K: The waveforms can actually be very complicated. You don't really see an earlier onset, and the first maximum in the checkerboard is late, not fast.

G: I bet you that means Diego has a bug in his code!

K: How much do you want to bet?

G: Last time I placed a bet with a young seismologist I lost a bottle of champagne, so I am getting a little more careful. But surely, you cannot cross-correlate the bottom and the top seismogram? They don't look in the least alike... That would be as hopeless as trying to do a full waveform inversion!

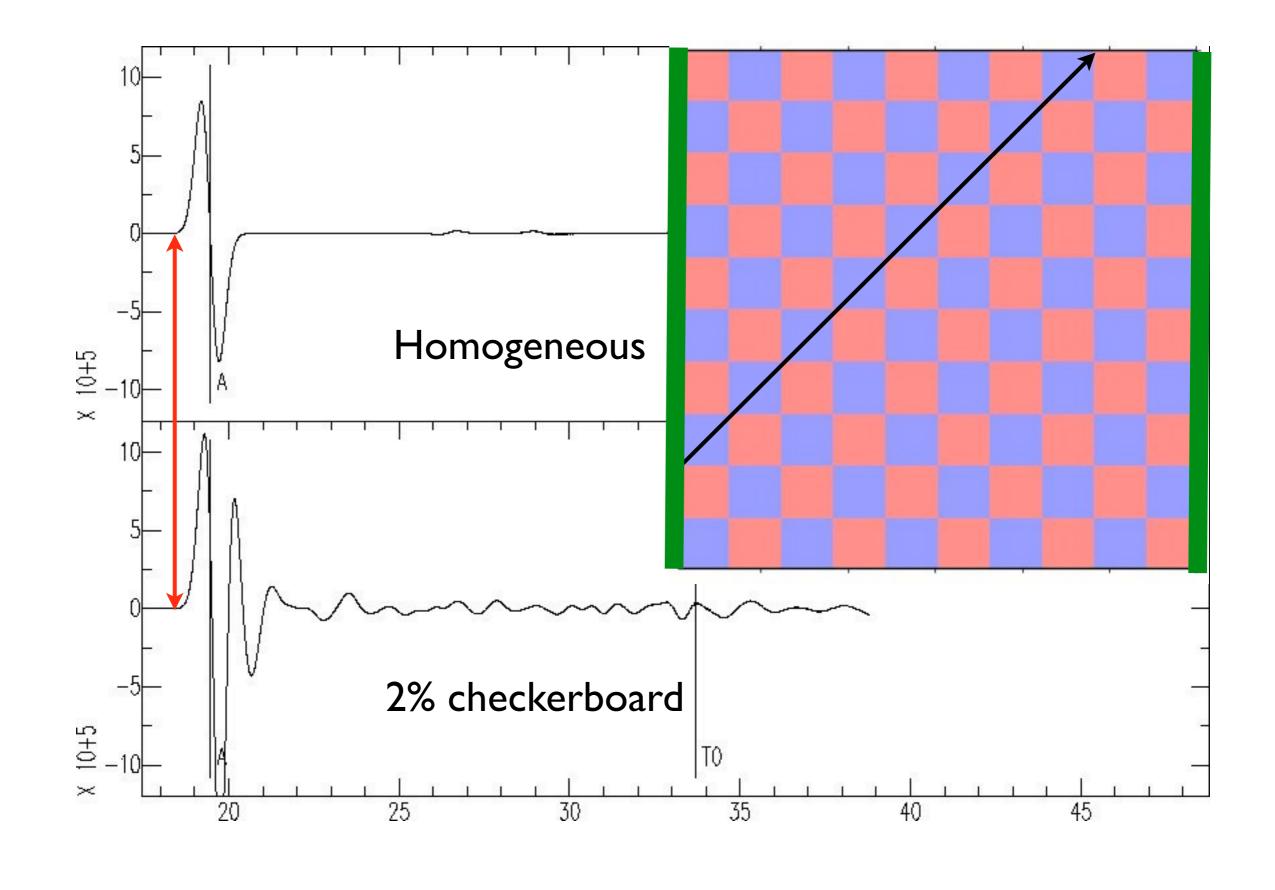


K: Well my correlation code does not say "no" if I tell it to correlate, so I can always do it. Here is the result.

G: Ha! But you can see it makes no sense,... the ray crossed fast fields only, and you measure a positive delay, so your cross-correlation thinks it has slowed down!

K: So what?

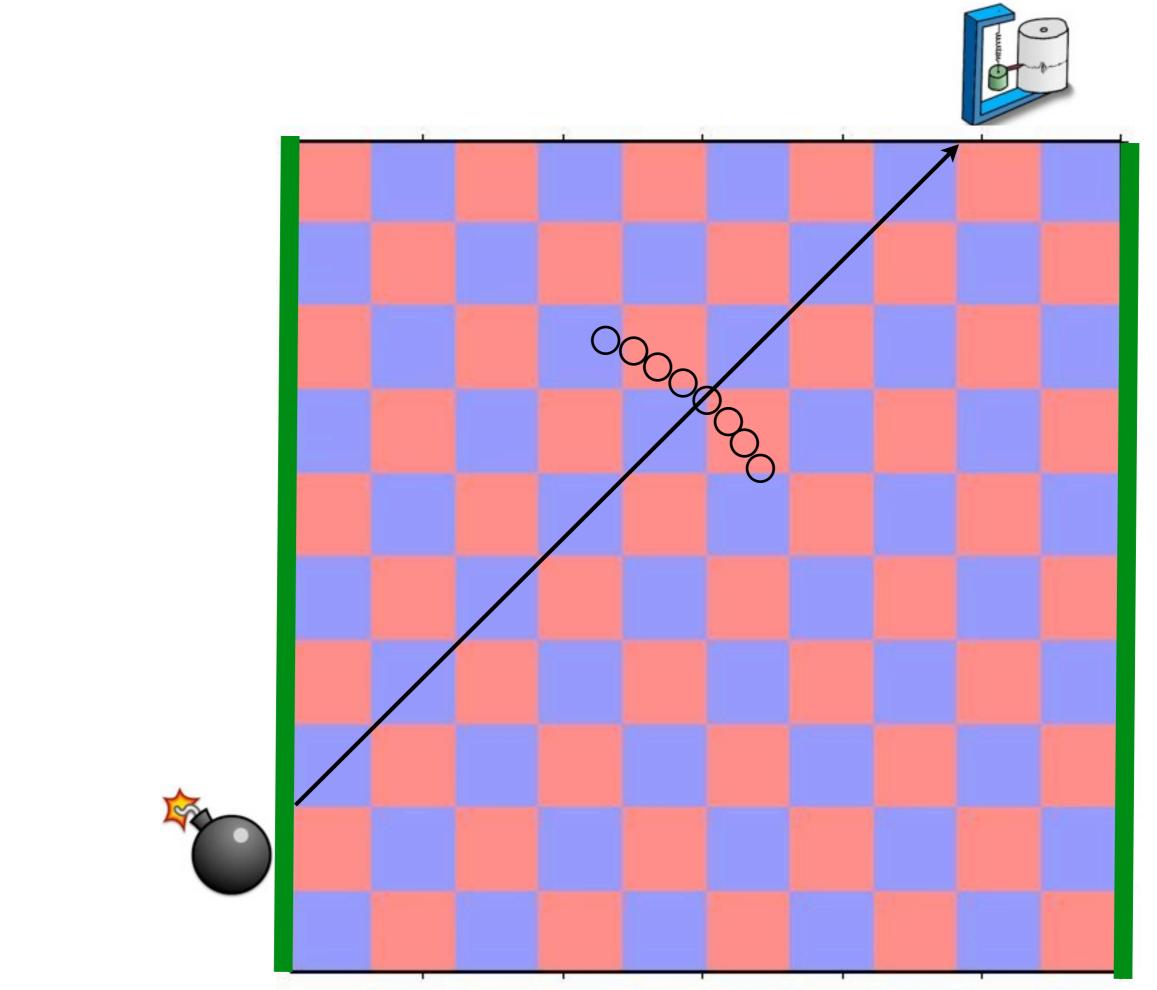
G: [acts dumbfounded]. Well the cross-correlation is obviously wrong, I told you so.



K: Not necessarily – the point is you think in terms of ray theory. But you cannot see the fast onset predicted by ray theory, so ray theory is useless.

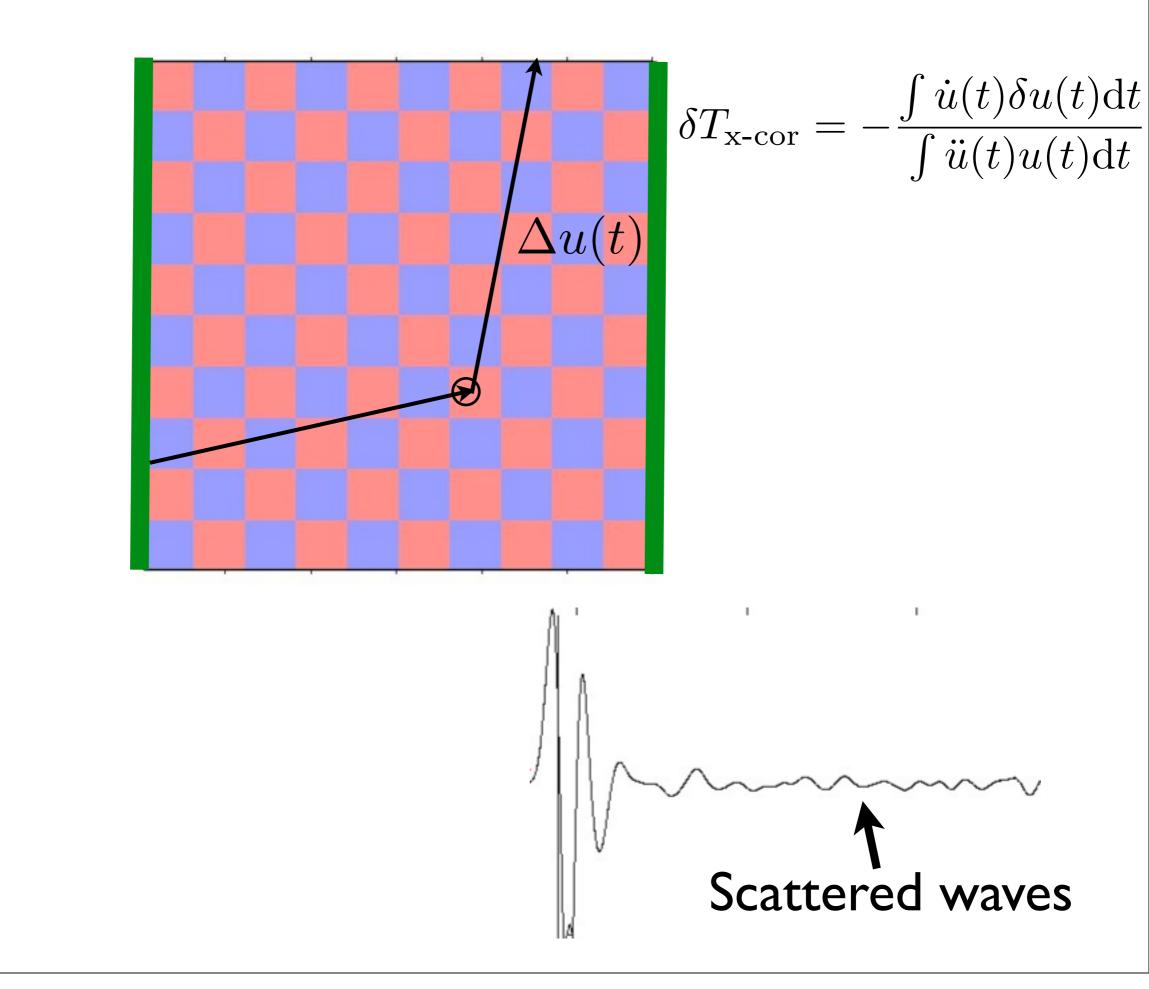
G: Mmm. I have to agree with that. But I'd still say we face a problem if neither ray theory nor cross-correlation agrees with my intuition.

K: Well, let us go back to Huygens



K: so think again of secondary sources on the wavefront. Huygen's Principle says that these all contribute to the signal and sample a lot of the slow boxes.

G: Yes – the modern version is the Kirchhoff integral, and you'll find that these sources are only interfering constructively for paths near the ray, which is why ray theory works so well. K: precisely; NEAR the ray – not ON the ray. The only path that senses no negative anomalies is the one path that crosses all the corners exactly. It has so little energy to contribute to the integral, that you cannot see its onset by eye – not even when there is no noise. The others sample both negative and positive anomalies, the net effect is small. G: small, but the wrong sign: not a speed–up as I would have expected.

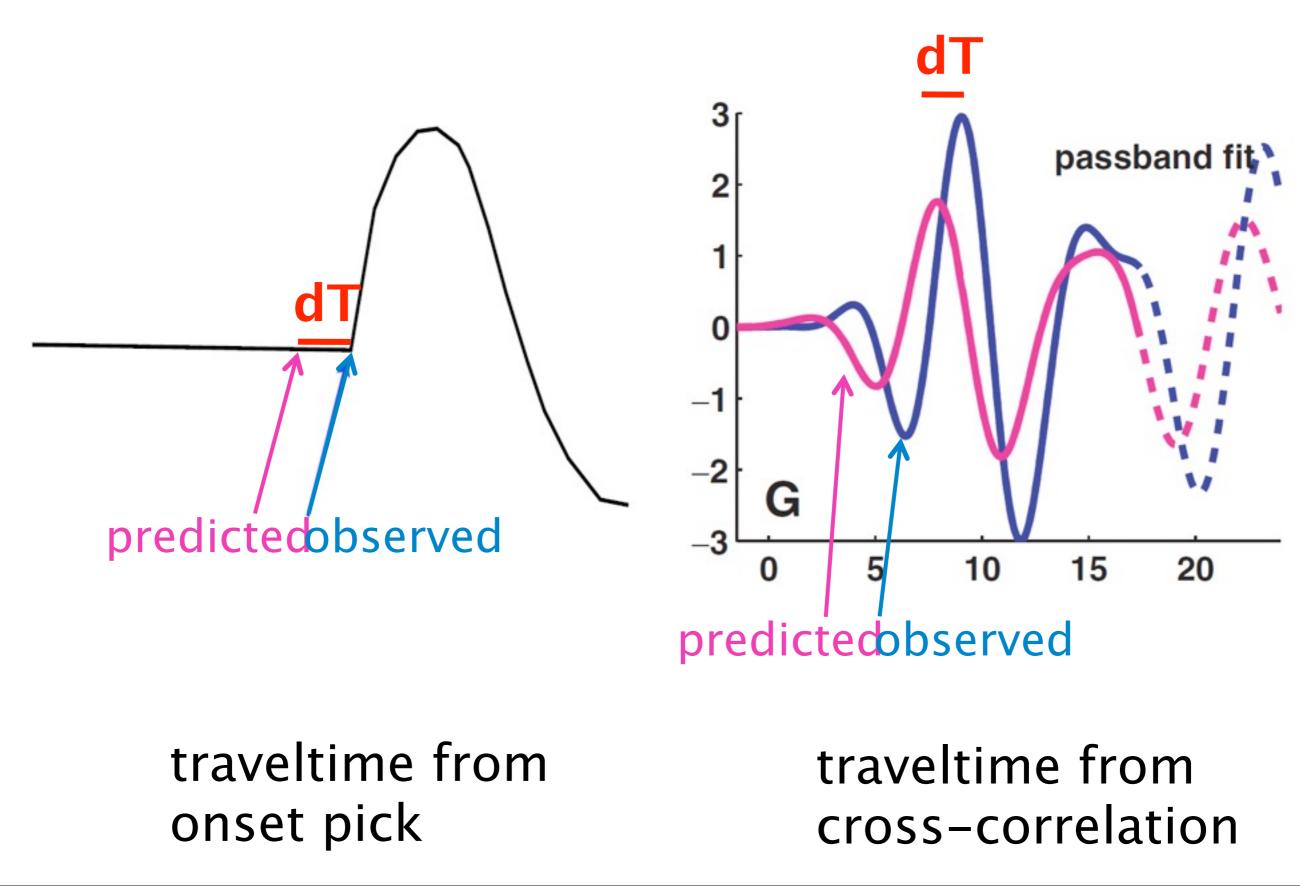


K: The sign is actually very difficult to predict. Remember this equation for the crosscorrelation delay?

G: Yes, it has the perturbed field delta u(t) in there, together with the unperturbed field. K: Correct. In our case the unperturbed field is if there is no checkerboard. The 2% anomalies give rise to a scattered field 'delta u(t)'. They act as real sources, you can see the scattered waves by eye.

G: You mean there is a lot of multipathing, different waves arriving in the same time window? K: Yes.

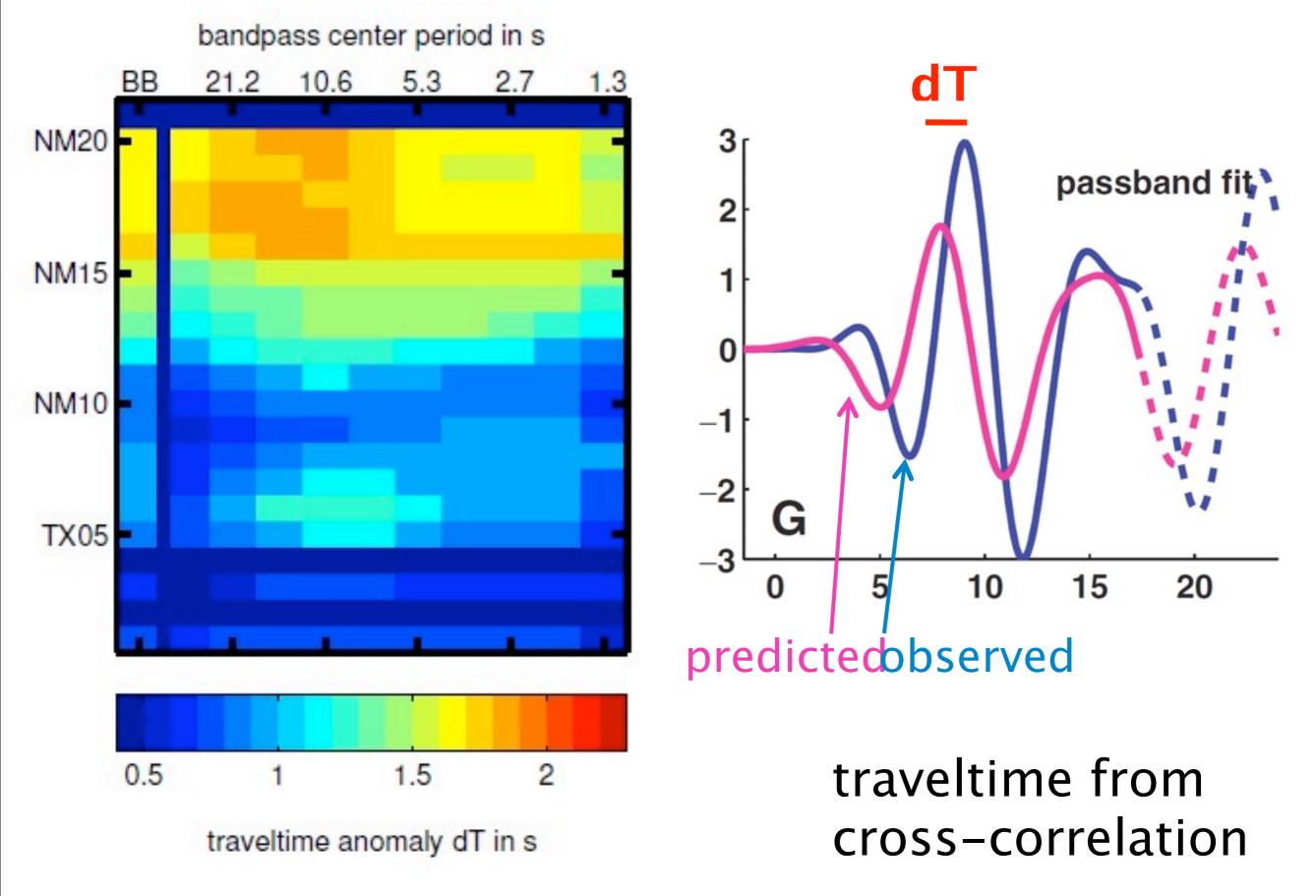
Fit every sample?



Monday, 21 May 2012

G: OK, but in the beginning you lectured me for using only one bit out of 1024. Really you are doing the same with your cross-correlation times. You have just replaced ray theory by more realistic modeling that includes scattering. But you are still using only one observation.

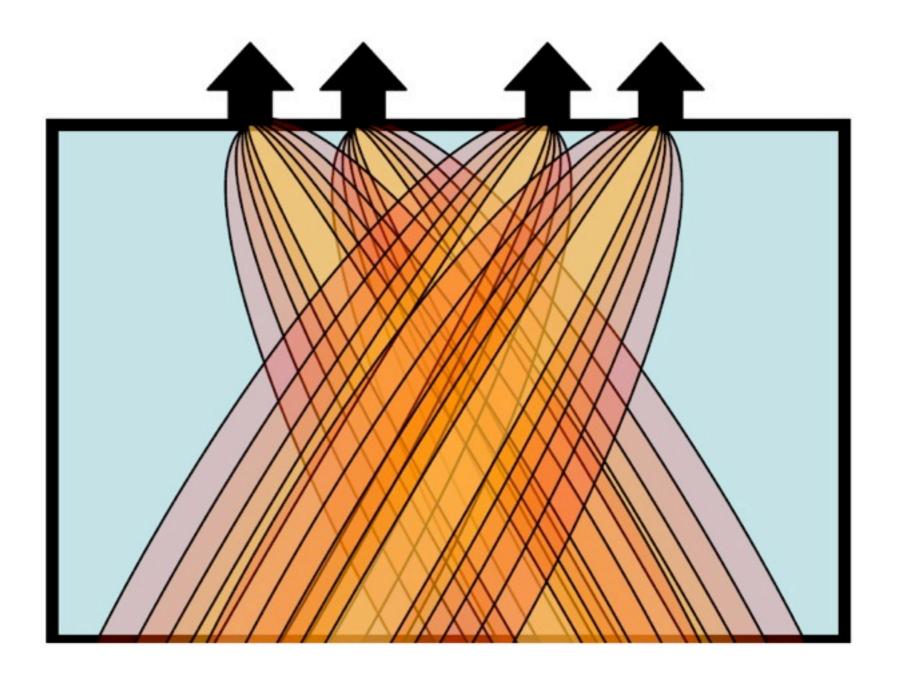
Fit every sample?



Monday, 21 May 2012

K: Not quite. We measure and model in different frequency bands. In these measurements of real P-waveforms there are 8 frequency bands.

Fit every sample?



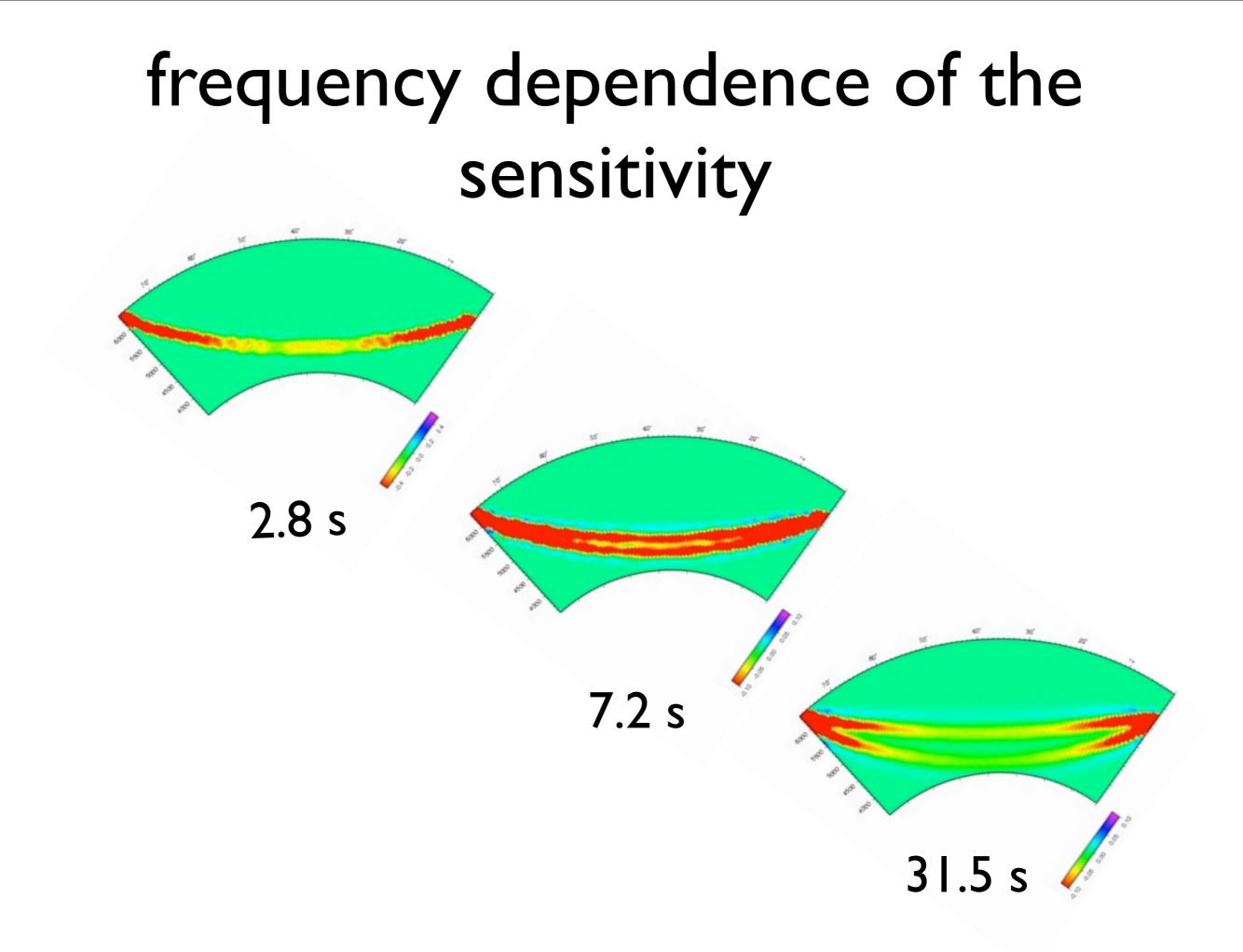
Multiple constraints from every wave path

Monday, 21 May 2012

K: The idea is that every wavelength "sees" the earth in a different way and thus gives a complementary piece of information. Each measurement is associated with a different sensitivity, and we force the solution to satisfy ALL of these independent constraints.

G: OK, the seismogram of length 1024 is condensed into eight scalars instead of one. But that still leaves 1016 degrees of freedom unused. Didn't waveform inversion promise to use all samples?

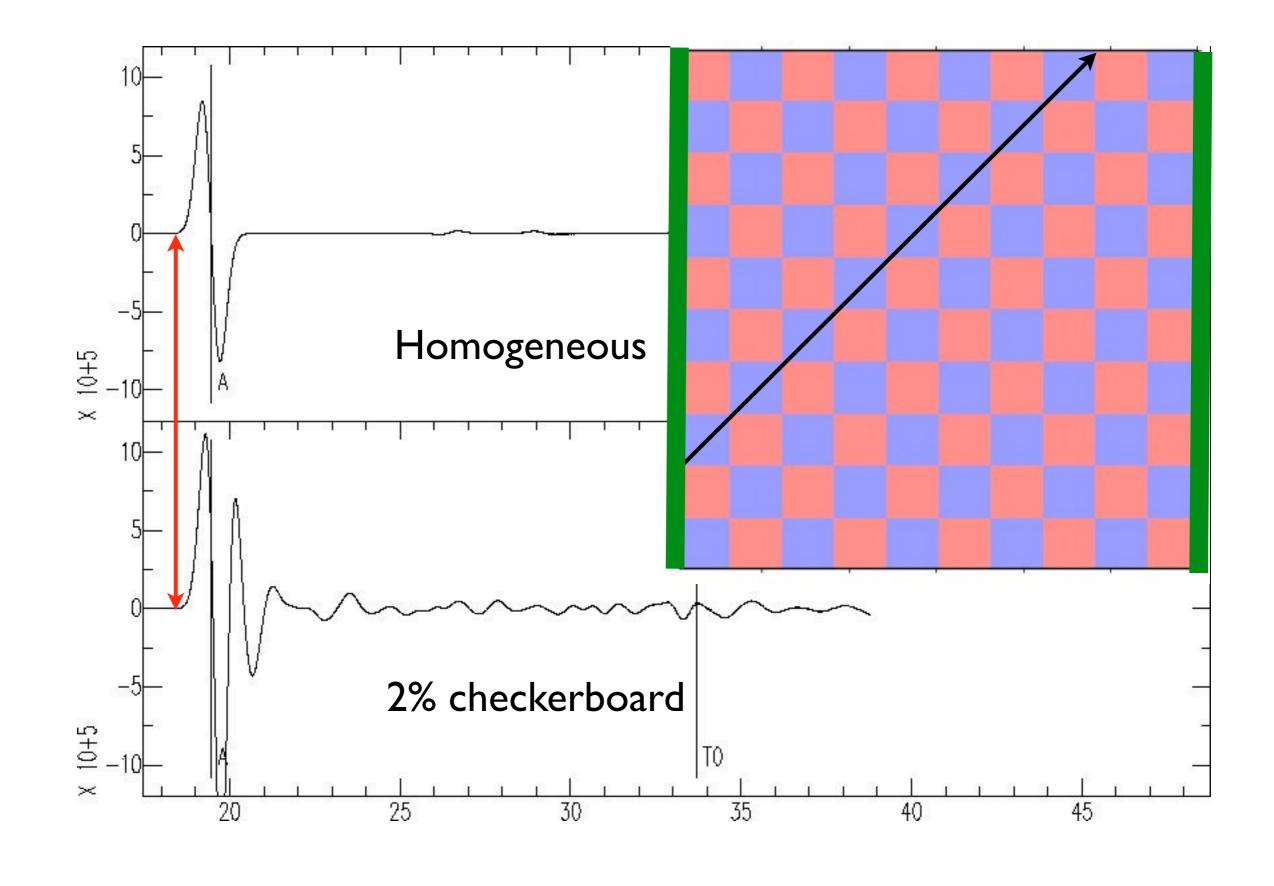
K: Yes but we get the most important information, the rest becomes more and more redundant.



K: Here is how it works on large scale in the earth, where one can easily measure bandpassed signals with periods from 2-30 seconds, each with their own delay.

G: So, we cannot just speak of one 'arrival time'?

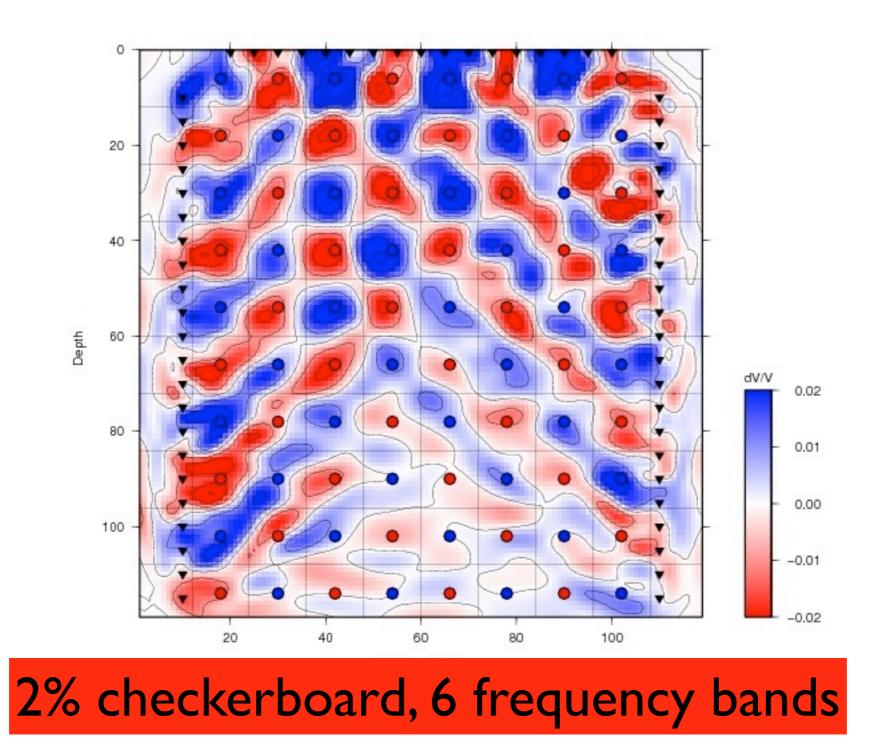
K: Only in the limit of zero period, if you wish. But it is better to acknowledge that the crosscorrelation delay is a different beast, for which we have to use a theory different from ray theory.



G: How is this dispersion going to help us?

K: Well, if you band-pass the signals at some low frequency, you are going to include the contribution of all the wiggles that arrive at later times. Those are the scattered waves that constitute delta u.

It works very well:

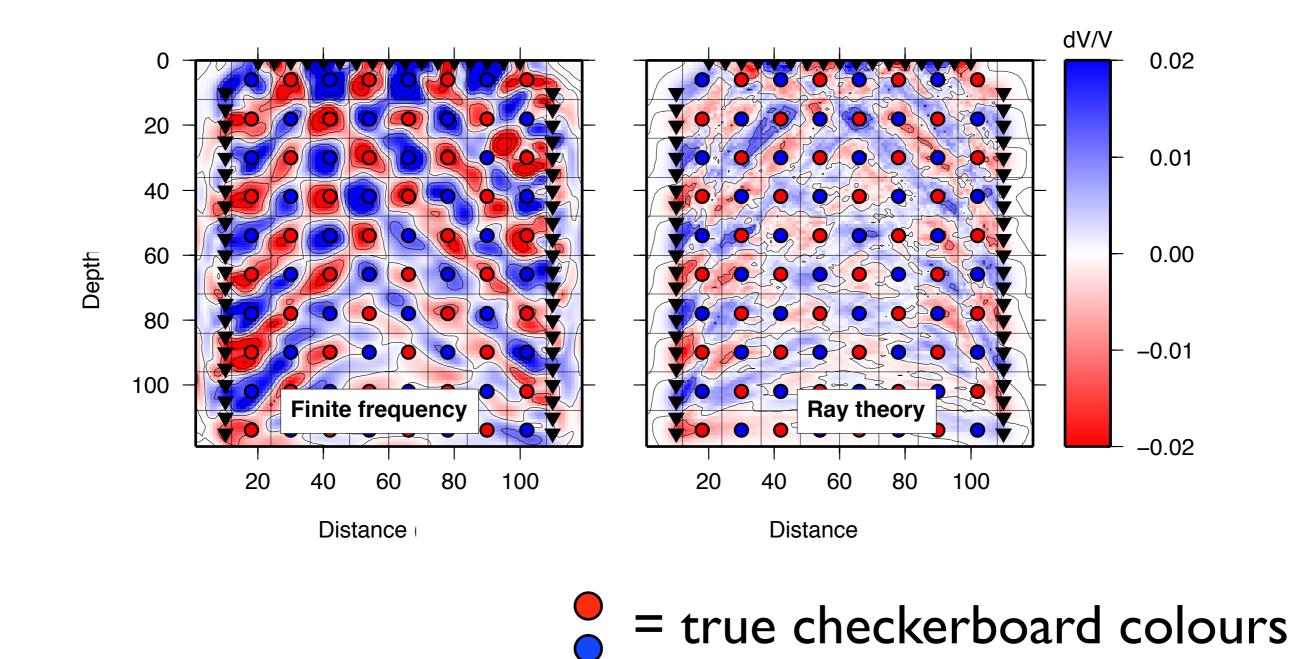


K: here is a solution of cross-correlation times that were actually measured from the synthetic seismograms. You can see the fields in the top half have very much the correct shape and amplitude. The little circles give the true anomaly amplitudes.

G: But could one not get the same result with ray theory?

K: You would have to measure onset times, and as we have seen, these are usually invisible. And if you use cross-correlation delays, ray theory is not the correct theory to interpret them.

The Devil's Checkerboard

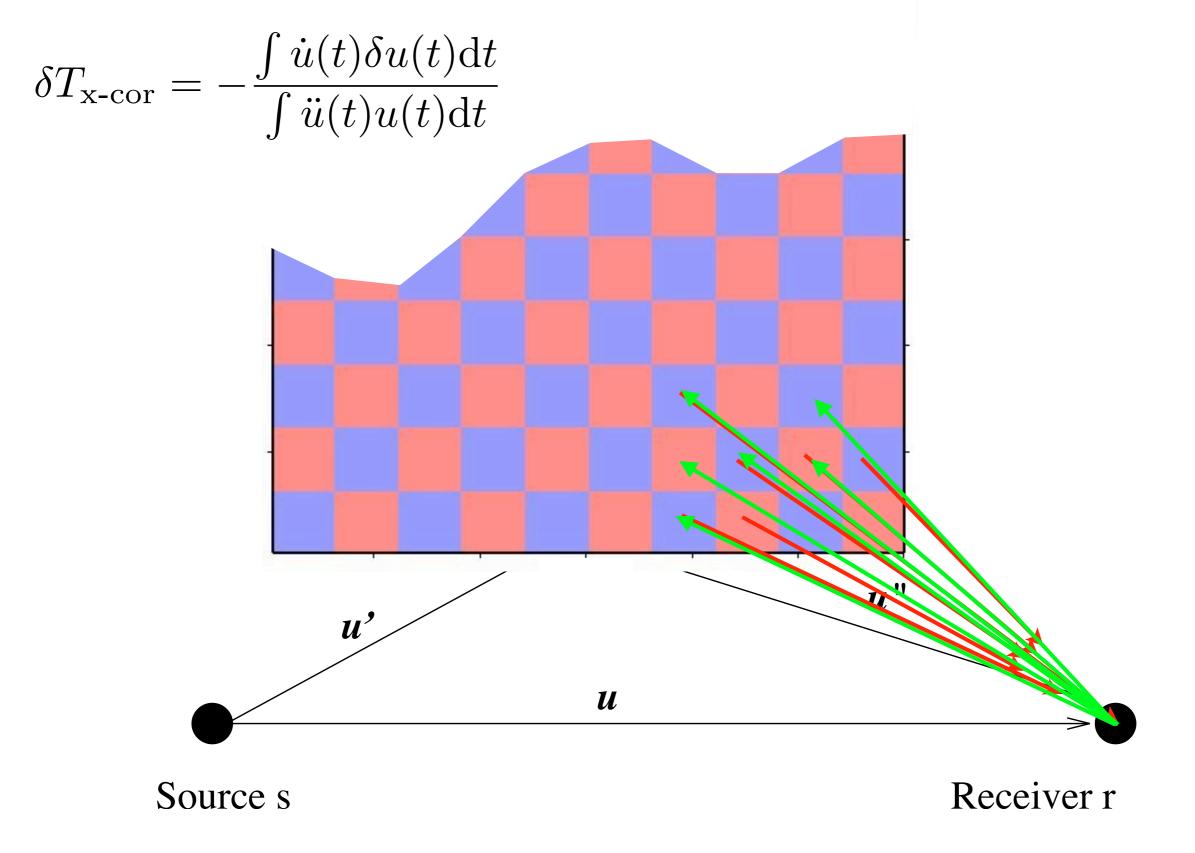


Monday, 21 May 2012

K: Diego repeated an experiment that Bo Jacobsen and I did some time ago, but this time with his treasure of synthetic seismograms in the 3D checkerboard and actual observed delay times. He took the measured delays in the 2 ms band and inverted them with ray theory. In the image on the right you can clearly see that this can lead to complete colour reversal in

- many regions of the checkerboard depending on the distance from the sources and receivers in the boreholes and at the surface.
- G: Wow. When I saw that the cross-correlation delay for that fast raypath was positive I concluded that cross-correlation does not work. But that positive delay gives the right checkerboard on the left! So finite-frequency theory is *really* different from ray theory! K: I'd rather say that ray theory is simply wrong in this case.

adjoint computation of δu



Monday, 21 May 2012

G: OK, I am beginning to understand. But I guess this is all a bit too late, isn't it. Everyone tells me the banana-doughnut kernels have been taken over by adjoint kernels? K: That is actually a very confusing terminology. What others call adjoint, we call reciprocity in ray theory. It is just a way of avoiding that one has to compute the wavefield from many

- Huygens sources. [click] you compute delta u from the receiver, and with reciprocity you can interpret this as the field from many different sources in the receiver.
- G: So why all the broohaha?
- K: because if you go beyond ray theory and you compute the complete wavefield, with reverberations and diffractions and the like, you have to do a second finite difference or spectral element computation, which is not trivial. With ray theory to compute the kernels, it is trivial.

Matrix solver or gradient search?

(Delay time data) $oldsymbol{A} oldsymbol{m} = oldsymbol{d}$

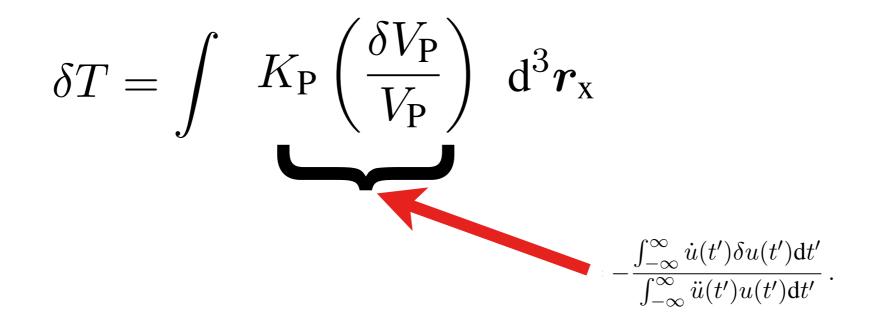
(Waveform data) $oldsymbol{Am} = oldsymbol{\delta u} \ oldsymbol{m} = (oldsymbol{A}^Toldsymbol{A})^{-1}oldsymbol{A}^Toldsymbol{\delta u} pprox lpha oldsymbol{A}^Toldsymbol{\delta u}$

Monday, 21 May 2012

G: Is that all?

K: Well, the adjoint method comes from full waveform inversion, in which the data are not delays, but the seismograms themselves. The vector d – and thus the matrix – is in that case huge. The matrix cannot be stored in memory and you are forced to search in a gradient direction. A simple representation of the difference is shown here, though it is a bit more complicated in practice. Once you have a linearized relationship with the seismogram perturbation, you take the adjoint (or the matrix transpose) to find an improvement to the model.

What about linearity of the kernels?



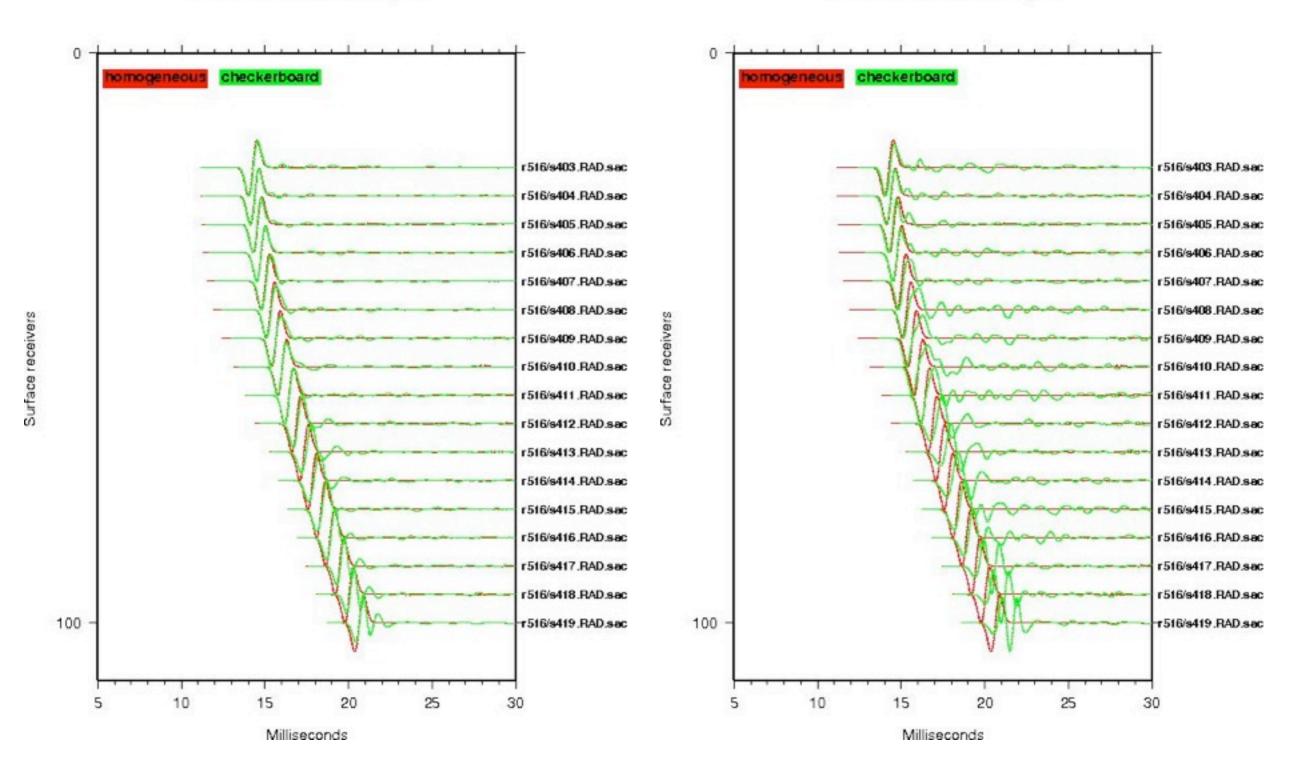
Monday, 21 May 2012

G: But I have been told adjoint inversions are linear and banana-donut inversions cannot be nonlinear?

K: That is mostly hype. Both are linearized. Just as one can re-compute wavefields in a second or third iteration of an adjoint inversion, you can recompute rays. All it takes is a 3D ray-tracer. The adjoint method will eventually include second order scattering into account, but the question is if that is important.

Source r516 checker2pct

Source r516 checker5pct

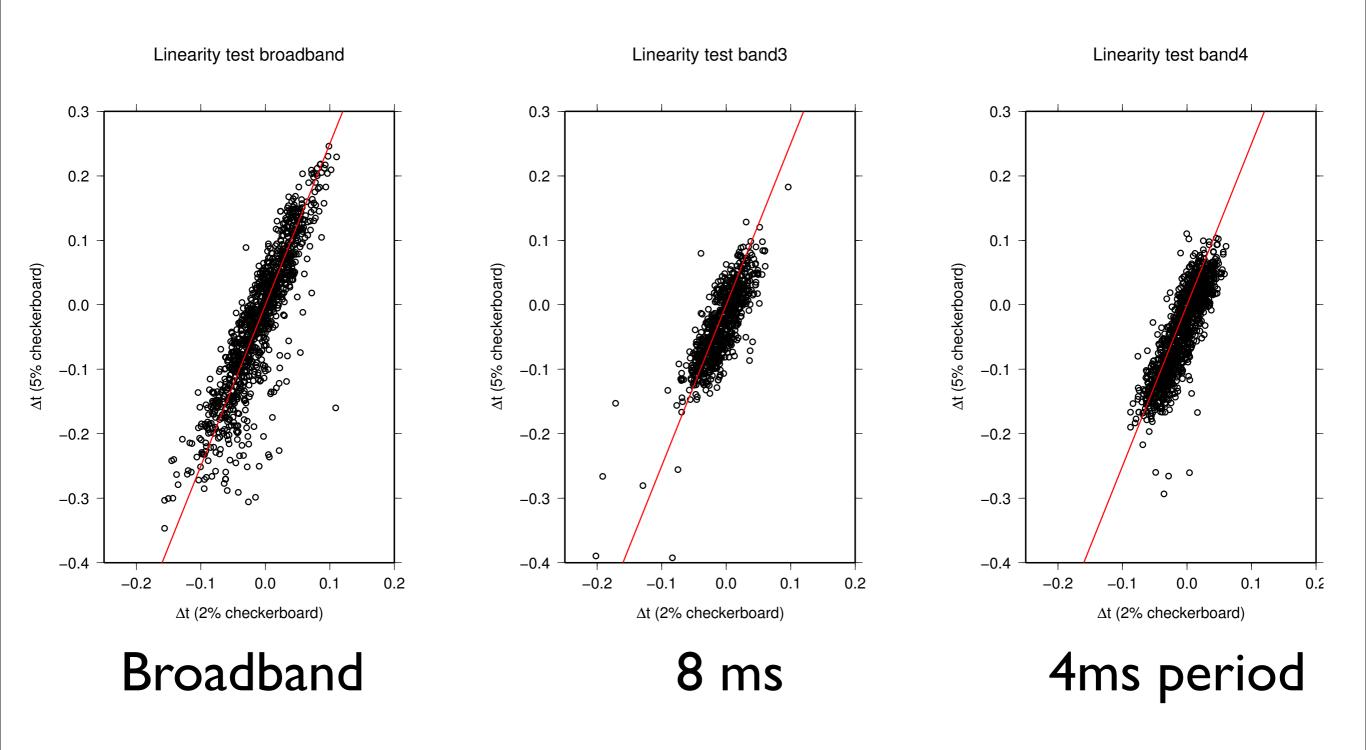


Monday, 21 May 2012

K: To check on how linear this is, Diego redid the checkerboard test with 5% anomalies. G: I guess the seismograms on the right are for the stronger anomalies? They look more complex.

K: Correct. First of all, if the linearized expression for the cross-correlation is correct, we expect the delay times to have a ratio of 2 to 5.

Linearity of ΔT

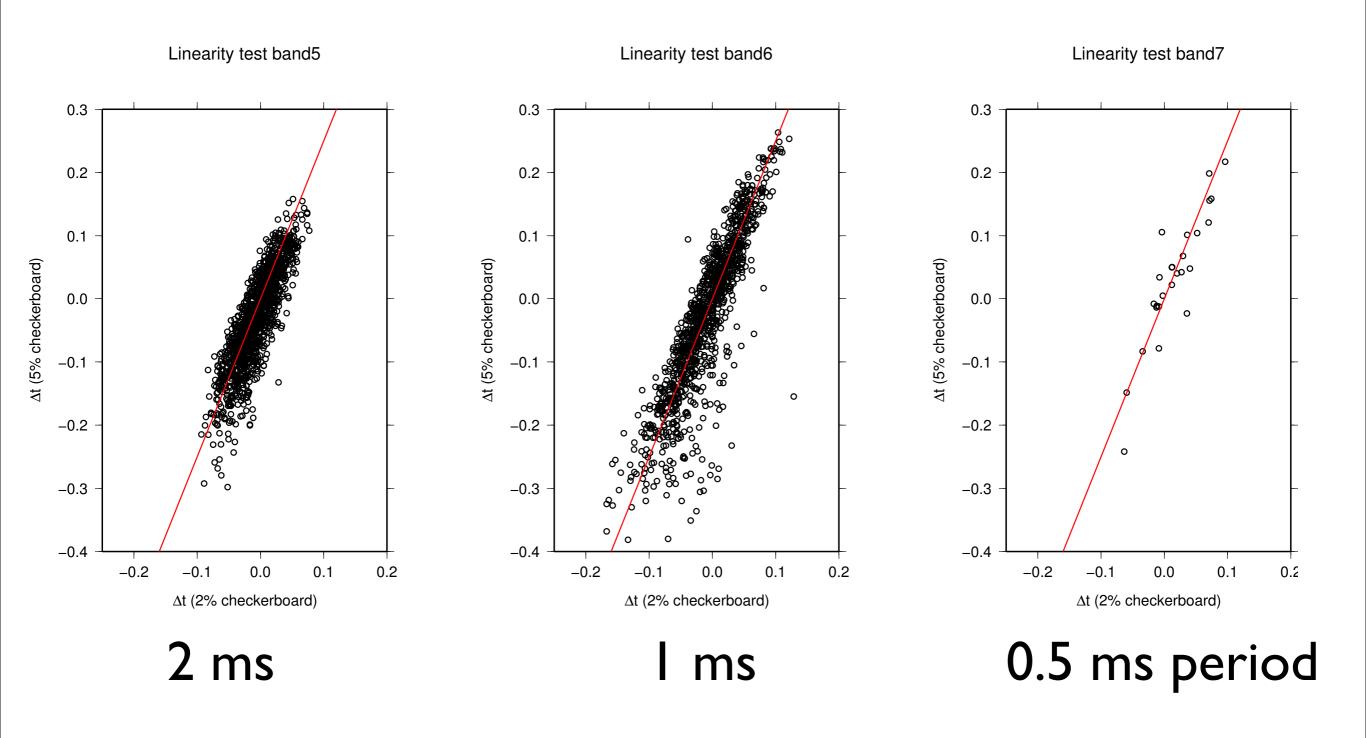


Monday, 21 May 2012

K: When we compare the cross-correlation delays of the checkerboard with 5% anomalies with those for the 2% case, we see that they are scattered because of errors in picking the maximum, but they are in the ratio of 2:5. So at least for anomalies of up to 5% we remain in the linear domain.

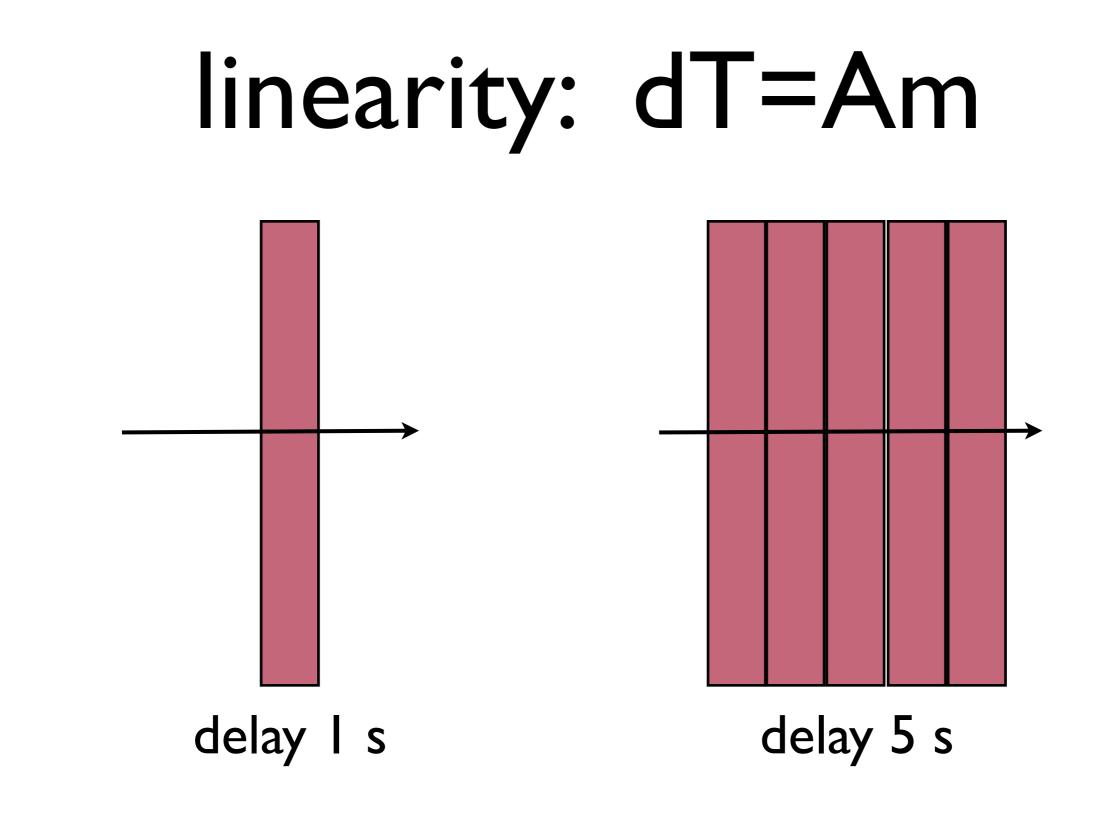
G: But hold it – why are there so much fewer data at bandpassed records than in the full frequency band at the left?

- K: Can't you guess?
- G: [silent to let audience think]



K: actually, at the higher frequencies, the spread of anomalies is much larger. G: Ah, I get it! The low frequencies are more concentrated towards zero...The low frequencies suffer more from the wavefront healing. Instead of going along a red or blue diagonal line they pick up the other colours of the neighbouring checker cubes...

K: Yes. That was actually causing the disastrous effects in the devil's checkerboard..



Monday, 21 May 2012

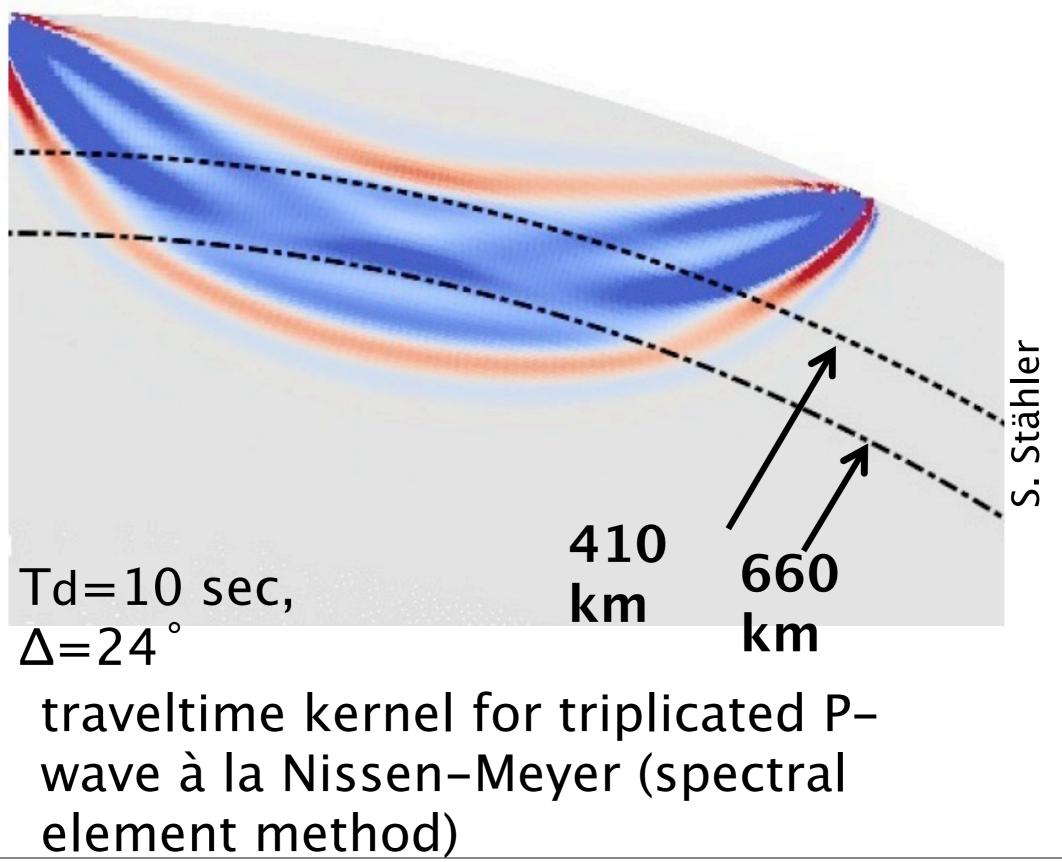
K: we have thus a practical proof that cross-correlation delays scale linearly with the model. Actually, that is what one would expect if ray theory is valid: a wide wall that is five times thicker will give a delay that is 5 times larger. We now know that it is also valid if ray theory is not OK, at least up to 5% – which is enough for the mantle of the Earth. G: and the theoretical relationship is also linear! The banana-donut kernels would give a five times larger delay on the right. K: precisely. That means that BD-kernels can handle large delays, even if they are larger than a quarter period of the wave.

full waveform inversion? linearity of du

Monday, 21 May 2012

K: But the same is not true for waveforms! Even though the delay delta T depends quite linearly on the model, the linearization of harmonic functions is quite troublesome.

Keep fitting phases / traveltimes?

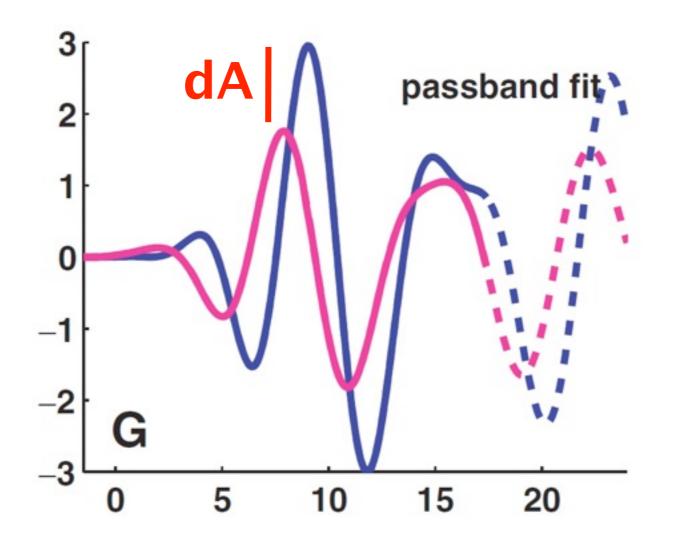


Monday, 21 May 2012

K: Travel times, which are essentially phase shifts in the passbands, are more linear and should thus be more stable. Especially given the levels of random and systematic noise in real data. We can ask the audience who actually has experience waveform inversion. I get the sense that in the "adjoint" approach as well, the

emphasis today is generally on fitting the phase first and foremost.

Also fit wave amplitudes?

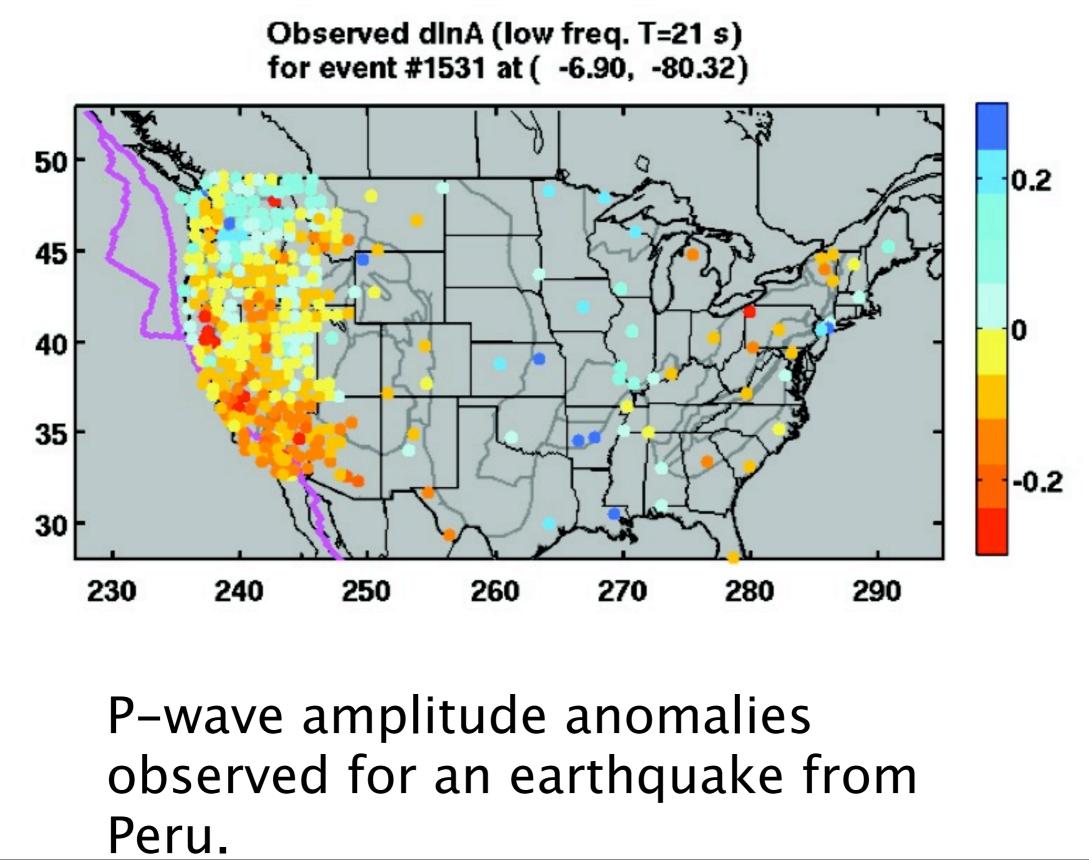


Monday, 21 May 2012

G: We have talked about delays. You have tried to fit amplitudes though? K: Yes, I tried to fit amplitude anomalies, as a second robust characteristic of the seismogram, besides traveltimes. Amplitudes have very different measurement sensitivities from traveltimes.

G: The motivation was to estimate attenuation?

Also fit wave amplitudes?

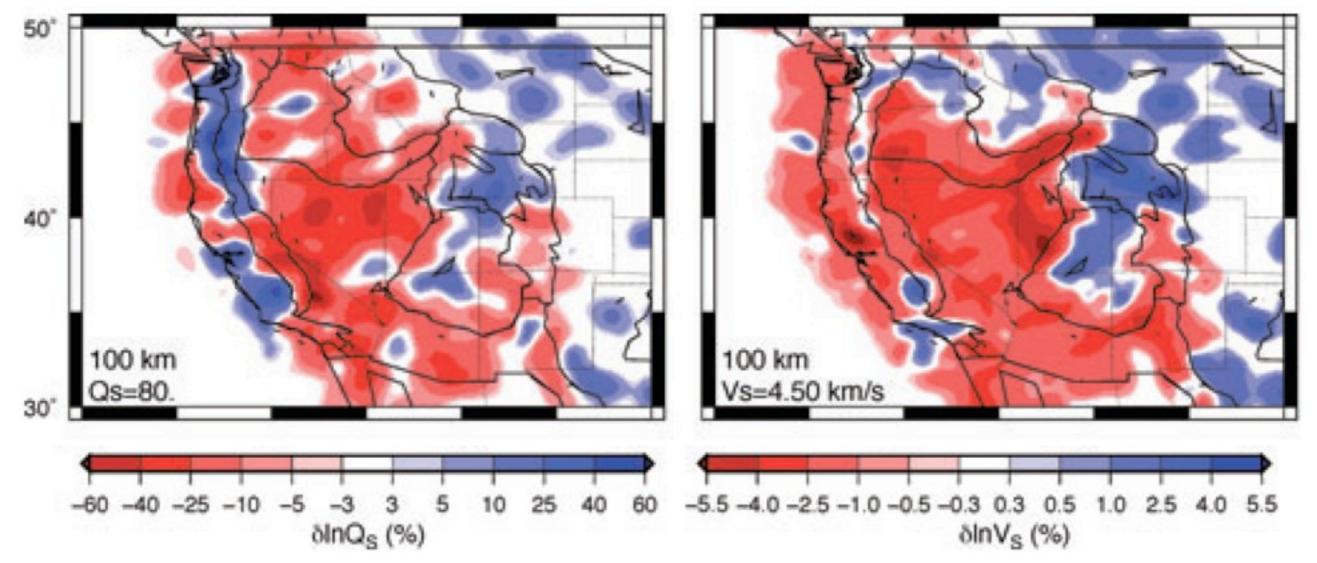


Monday, 21 May 2012

K: Yes, but only after correcting for elastic effects (focusing). Most of the observed signal here is due to focusing, we found.

Also fit wave amplitudes?

Tian et al., 2009



Attenuation structure from inversion of Swave amplitude anomalies for dlnQs and dlnVs jointly.

Monday, 21 May 2012

K: Getting at Q sort of worked for S-waves, underneath the dense USArray. But for Pwaves, the anelastic signal drowned in the measurement noise, despite accounting for focusing.

Do we need waveform inversion when there is lots of

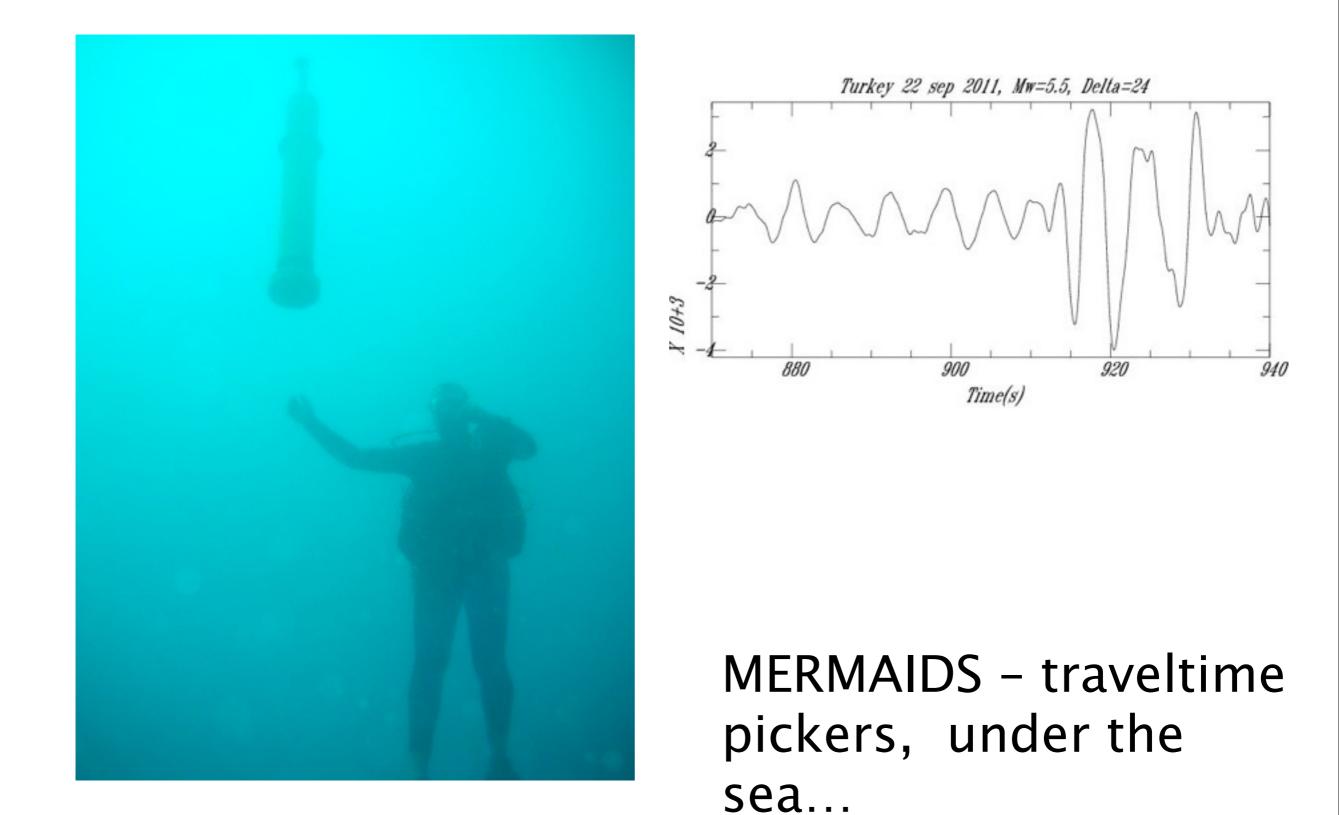
NOSE?

Monday, 21 May 2012

G: Welcome to the real world. So far we talked about theory, synthetics tests, and favorable geometries. But data are noisy. And we never have enough data. How useful are those sophisticated waveform inversions, when data quality and coverage are poor? K: That's a whole new lecture. I don't think this question has seen much effort at quantitative answers. NOISE, we can think about that for next year. You won't have retired, will you?

G: No, not before I release my MERMAIDS to the ocean.

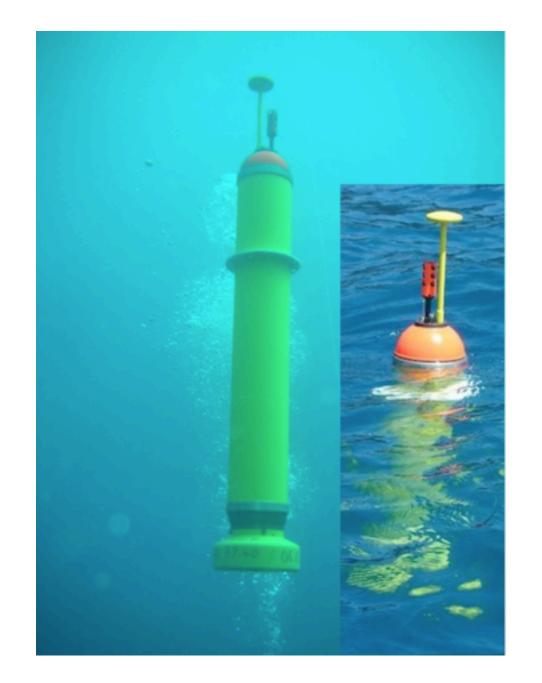
Waveform inversion and noise?

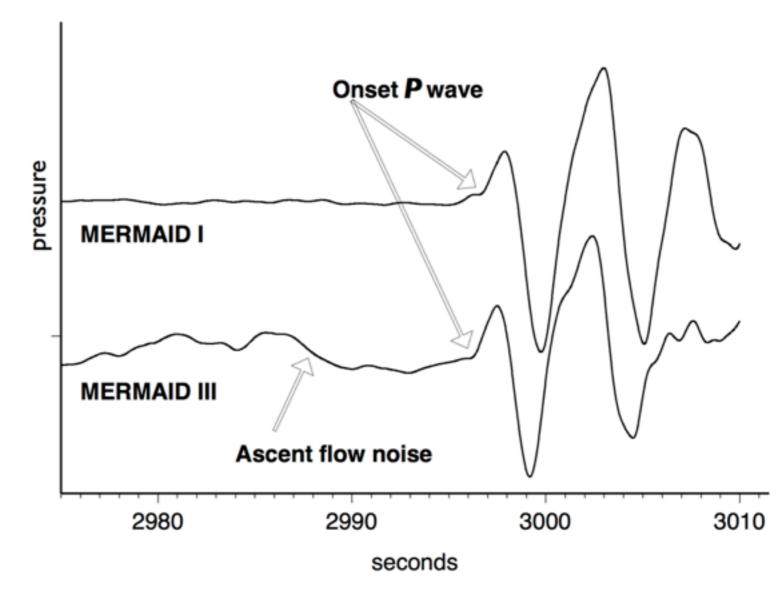


Monday, 21 May 2012

K: What do they do?

G: They float in the oceans at around 1000 m depth. When they detect an earthquake, they come to the surface and send a short piece of seismogram via satellite. Even if we can only see the onset we'll have completely new information. So far, we have almost no data from the oceans. K: Does that mean back to ray theoretical inversion again?





Fox Islands, M=7.4, Δ =85

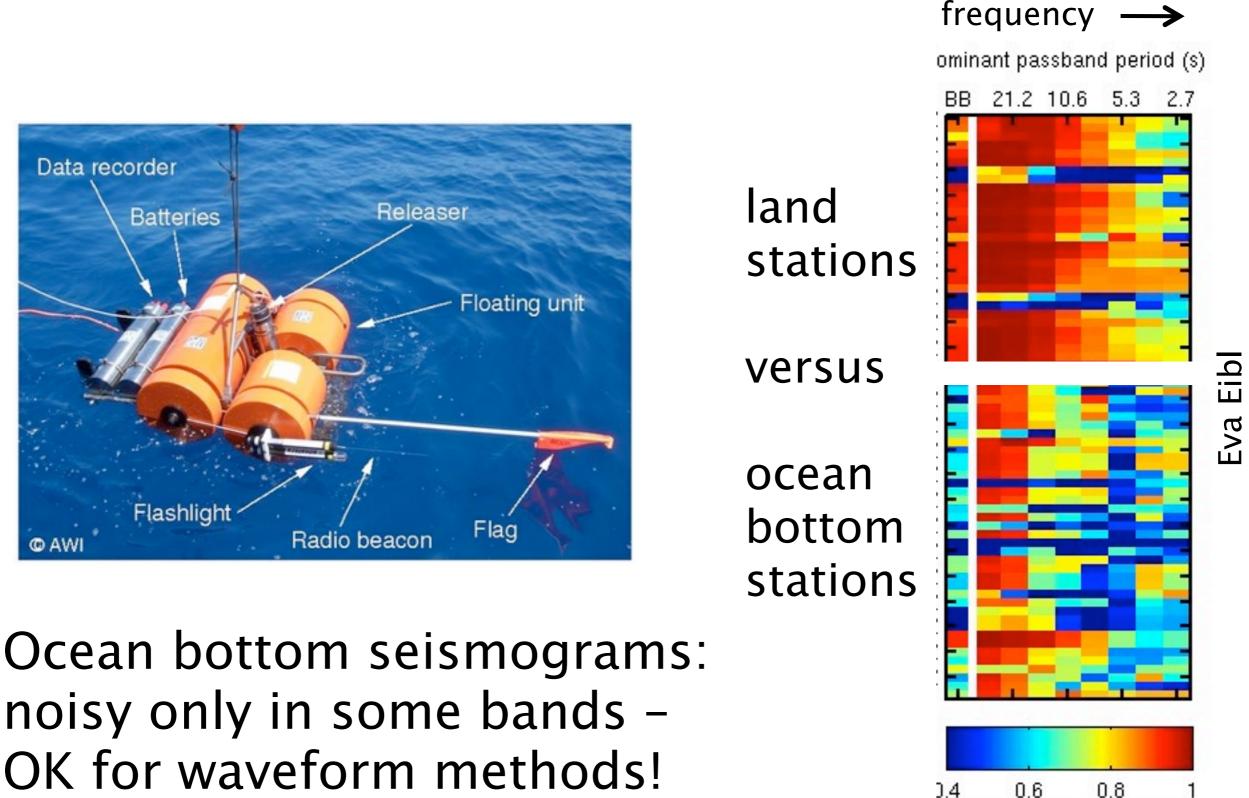
45

Monday, 21 May 2012

G: In some cases – when all we can see is the onset, yes. But for stronger signals like this Fox Island quake of magnitude 7.4, we can not only pick the onset but also cross-correlate P waves across a network of Mermaids, waveforms are in this case sufficiently similar. In october we shall launch half a dozen Mermaids near La

Reunion, in your Rhum–Rum experiment!

Waveform inversion and noise?



I.4 0.6 0.8 1 x-correlation coefficient

Monday, 21 May 2012

K: On that cruise, we will also deploy ocean bottom seismometers. They also have a reputation for being noisy. The reality is more mixed. Some frequency bands are so noisy as to be useless. Others have a very good signal quality. It's a non-broadband situation where it is impossible to pick an onset. But for finite-frequency methods, that does not pose a problem. On the contrary, such data can be accommodated naturally.

Conclusions I

- Noisy data often require correlation
- Cross-correlation delays require finitefrequency theory

Monday, 21 May 2012

G: So if I have understood you correctly, you agree that x-correlations are more precise, K: Yes, but if you do that you must use finite frequency theory or you're in trouble

G: Like with the checkerboard inversion, I got it...

Conclusions II

- With respect to waveform inversion:
- They remain linear for earth-like anomalies
- They reduce the matrix size, gradient search can often be avoided
- For body wave signals: ray theory offers 2-3 orders of magnitude speed-up in computation

- the kernels are by definition linear and the checkerboard tests clearly show that the delay times follow suit. Since they agree for small anomalies they do too for the large ones. G: So one does not really need to use full waveforms?
- K: Once you've split the time series in windows with their frequency-dependent delays, there is little extra value in milking the seismogram even further. I'd rather use more different source-station pairs. If the windows contain bod waves, ray theory will also speed up the matrix computation but the windows can in principle contain anything.

Monday, 21 May 2012

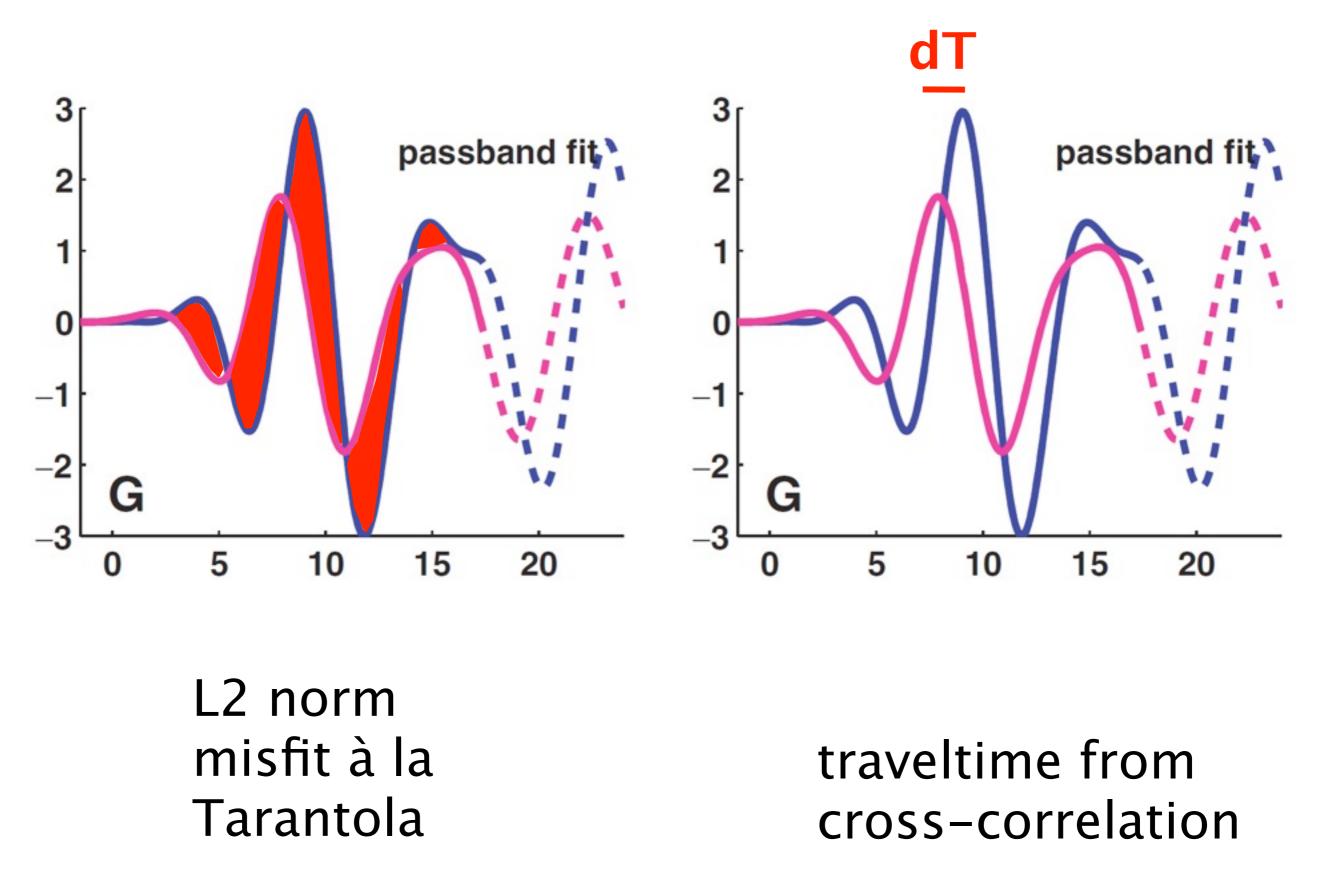
G: The fact that they remain linear even for complicated anomalies of several percent, was that well-known?

K: Not exactly – there was a discussion whether the linearity of BD-kernels was warranted, often linked to the linearity of the Born approximation. But that is the wrong way to look at it:

Thanks to....

- Diego Mercerat, now at Laboratoire Régional des Ponts et Chausées (CETE), Nice, France
- European Research Council
- QUEST
- All those anonymous reviewers who inspired this discussion with their stubborn judgements based on ray theory....

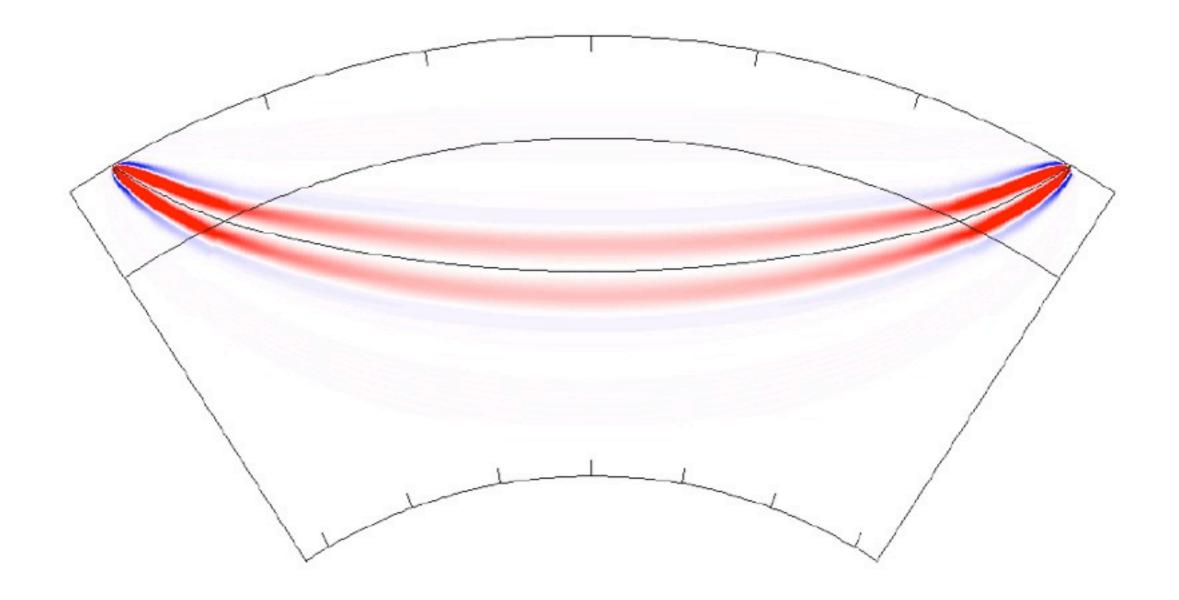
Keep fitting phases / traveltimes?



Monday, 21 May 2012

K: Sort of. In its original version proposed by Tarantola, the misfit criterion was the L2 norm of seismogram-minus-synthetic. So you explicitly weigh all samples evenly. But in the cross-correlation measure all samples influence the observed delay as well, we just do not overdo it.

Keep fitting phases / traveltimes?



traveltime kernel for P-wave à la Dahlen and Nolet (from paraxial ray tracing)

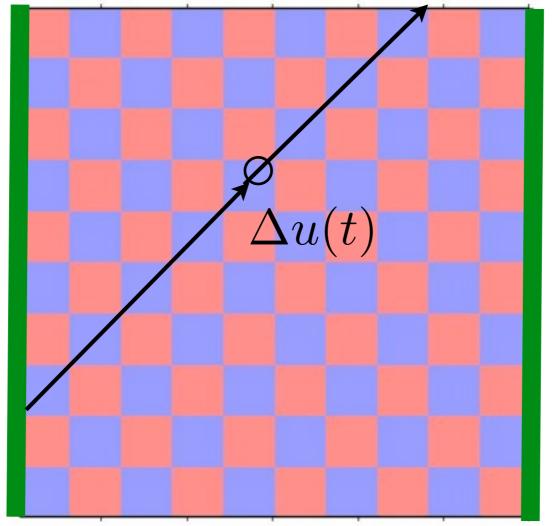
Monday, 21 May 2012

G: That brings us back to our earlier discussion on linearity. You are saying you don't want the raw L2 misfit because it is not linear.

K: Well, with the original finite frequency modeling a la Dahlen and Nolet, it would not have been possible to compute kernels for the raw L2 waveform misfit. You

know that, of course.

G: Yes, that limitation was due to the efficient approximation in which the kernels were computed from paraxial ray tracing. But you are now synthesizing kernels from full forward wave propagation a la Nissen-Meyer. And you still don't use the raw L2 norm?



$$\delta T_{\text{x-cor}} = -\frac{\int \dot{u}(t)\delta u(t)dt}{\int \ddot{u}(t)u(t)dt}$$

Monday, 21 May 2012

K: Yes. But now look at the contribution from a scatterer on the direct ray path. This will arrive at the same time as the direct wave. It can add to its amplitude, but it will not perturb its phase. And the result is that it does not affect the cross-correlation time.

G: You mean an anomaly on the raypath only affects the amplitude, but does not delay the wave? But what about that one ray that goes through all the corners?K: Zero, nothing. It drowns in the contribution of neighbouring paths.G: Unless the frequency is infinite...

K: When is the last time you've seen a seismic wave with infinite frequency?