J. Geophys. 43, 235–242, 1977

Journal of Geophysics

Three Dimensional Seismic Velocity Anomalies in the Lithosphere

Method and Summary of Results

Keiiti Aki

Department of Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Mass. 02139, USA

Abstract. This paper discusses cruxes of the method for inverting the P-time residual data introduced by Aki et al. (1976a) and summarizes the results obtained by the method on 3-dimensional seismic velocity anomalies in the lithosphere under several seismic arrays around the world.

The velocity anomalies at shallow depths correlate well with geologic features in young, active areas such as California, Hawaii and Yellowstone, but the correlation is not apparent in old, stable areas such as eastern Montana and Norway. Significant small scale (20 ~ 50 km) lateral inhomogeneity is observed everywhere to the depth of 100 ~ 150 km, with the minimum estimate of root mean square fluctuation about 3%. The lithosphere-asthenosphere boundary seems to manifest itself as change in the roughness of anomaly pattern or in the trend of anomaly.

Key words: Lithosphere – Velocity anomalies – Arrays.

Introduction

Recently, Aki et al. (1976a) introduced a new inversion method for teleseismic travel time data obtained at a seismic array with some areal spread. The method is very simple, never the less it gives a 3-dimensional seismic image of the earth's interior. The method has been applied to five major existing arrays; NORSAR (Aki et al., 1976a), LASA (Aki et al., 1976b), Central California (Husebye et al., 1976; Zandt and Aki, 1976), Hawaii (Ellsworth and Koyanagi, 1976), and Yellowstone (Iyer, 1976). The purpose of this review paper is to summarize salient features of 3-D seismic images of the earth under these arrays, and compare the results with geology and tectonics of the array siting area.

Method and Assumptions

Let us first briefly describe our method for teleseismic data inversion, with special attention to the assumptions we make in our inversion method. We

start with an initial model, consisting of a stack of homogeneous layers with parallel interfaces overlying a "standard" earth. Our data are travel times, say t_{ij}, observed at the station of i for the event of j. Our initial model includes an assumption about the seismic source that the incident waves are plane waves with known direction of approach determined by the hypocenter locations given by NOAA.

For the initial model thus defined, we can calculate the ray path and travel time expected for each station-event pair. We then form the travel time residual, that is, the observed time minus the calculated time for the initial model. Typically, we use $2{,}000 \sim 3{,}000$ travel time residuals for each array. The root mean square of residuals are usually several times larger than the measurement error, which is roughly 0.1 s. So, we want to reduce the residual variance by perturbing the initial model. Ideally, we want to perturb the whole earth and earthquake locations. But the limited data do not allow this. We perturb the medium only to a certain depth immediately beneath the array, by dividing each layer in many blocks and assigning to each block an unknown parameter representing the perturbation of slowness in the block.

In an addition to the perturbation of slowness inside the block model, we allow one parameter for each event to be perturbed. That is the arrival time of the plane wave front at a reference point fixed to the array. This parameter can absorb the D.C. error common to all the stations, such as errors in origin time of earthquakes or errors in the standard travel-time table $t(\Delta)$. We don't, however, allow the perturbation in $\mathrm{d}t/\mathrm{d}\Delta$, or the slope of incident wave front.

A major restriction of our model is this fixed wavefront, which may be justified if the earth below our block model is laterally homogeneous. Evidence in support of this assumption is available from the data obtained in California as described later.

Another important restriction of our model is that we don't perturb the shape of interface between layers. This restriction, however, is not fundamental, because if we can reduce the size of block, the interface fluctuation can be simulated by the block model. Typically, we perturbed the velocity in about 300 blocks for each array. A small perturbation in a model parameter may be linearly related with small observed residuals. We have as many linear equations as the number of observed residuals, which can be put into the following form.

$$d = Gm \tag{1}$$

where the elements of d, G, and m are respectively

$$(t_{ij}^{\mathrm{obs}} - t_{ij}^{\mathrm{cal}}) - \overline{(t_{ij}^{\mathrm{obs}} - t_{ij}^{\mathrm{cal}})}^{(i)}, \quad g_{ijk} - \bar{g}_{ijk}^{(i)}, \text{ and } m_k.$$

t_{ij} is the travel time for the station of i and the event of j. obs refers to the observed value and cal refers to the value calculated using the initial model defined earlier. g_{ijk} is the travel time spent by the ray (ij) in the block of k. We sample only one block from each layer, in which the ray spends most of its time. The bar suffixed with (i) indicates the average over all stations

for each event. The subtraction of this average is a consequence of allowing the D.C. error mentioned earlier. m_k is fractional perturbation of slowness for the block of k. The equation for m_k is decoupled from the equation for the D.C. error.

The usual least squares solution for m_k is non-unique, because the matrix G has zero eigen-values, at least as many as the number of layers in the initial model. This is because the uniform perturbation of all the blocks in a layer does not affect observable d, and is, therefore, undetermined by our data. Additional zero eigen-values are possible if two blocks always share common rays. To solve this problem, we tried two methods: one is the generalized inverse or minimum solution, the other is the damped least squares solution in which the effect of small eigen-values is smoothed out. The generalized inverse gives a rough picture, which suffers from random error effect on eigen-vectors with small eigen-values. The damped least squares solution gives a smooth picture, with a loss of resolution in vertical direction as explained below. We applied the two methods to data from several arrays, and concluded that both give essentially the same image and the difference can be explained by difference in the smoothing kernel, or resolution matrix.

Following Lanczos (1961), we decompose matrix G using eigen-vectors with non-zero eigen-values.

$$G = U_p \Lambda_p V_p^T \tag{2}$$

where Λ_p is a diagonal matrix, with elements equal to non-zero eigen values. U_p and V_p are matrices whose column vectors are the corresponding eigen-vectors in data space and model space, respectively. The generalized inverse solution m^G can be written as

$$m^G = V_p \Lambda_p^{-1} U_p^T d \tag{3}$$

and the damped least squares solution as

$$m^D = V_p \frac{\Lambda_p}{\Lambda_p^2 + \theta^2} U_p^T d \tag{4}$$

m^D can be obtained by solving the following equation.

$$(G^T G + \theta^2 I)\, m^D = G^T d \tag{5}$$

Equation (4) shows that m^D is an approximate generalized inverse, in which the eigen-vectors with eigen-values less than θ are suppressed. m^D is also a special case of stochastic inverse (Franklin, 1970), where θ^2 is the variance ratio of noise in the data to fluctuation in the model.

Eigen-Vectors with the Largest and Smallest Eigen-Values

Let us take the case of California data studied by Husebye et al. (1976) and show the eigen-vectors for large and small eigen-values. Twenty six stations roughly uniformly distributed in a rectangular area about 100×200 km are used in this study. The travel times were read for 66 earthquakes with fair azimuthal and distance coverage. The total number of time residual used in the inversion is 1517. The initial model consists of 5 layers with thinkness 25 km, each divided into 9×9 square blocks with side length 25 km for the

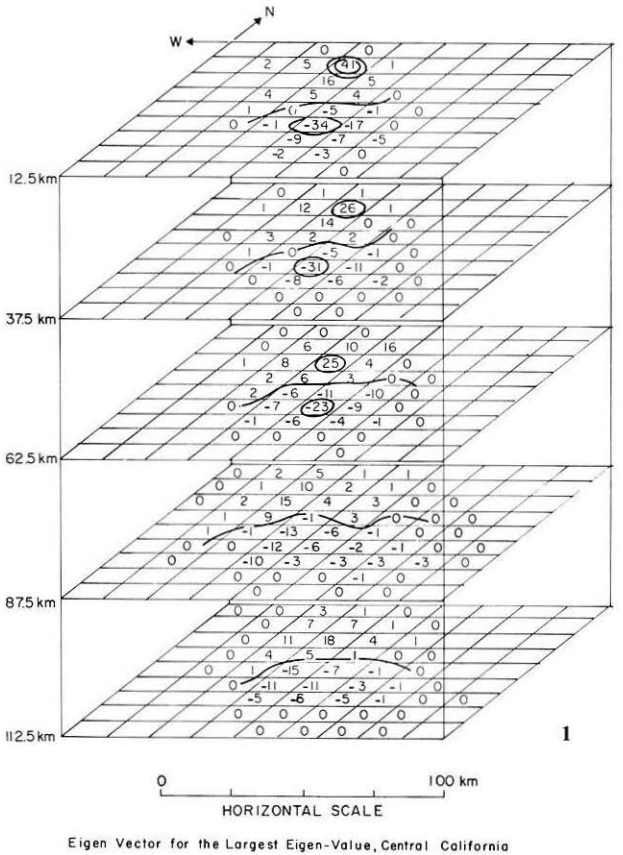

Fig. 1. Eigenvector for the largest eigenvalue in the case of central California studied by Husebye et al. (1976). Each sheet represents a layer in the initial model, placed at its median depth. The number in each square represents component of the eigenvector for the block in %. The contours are drawn at 20% interval. Although the block size is shown to have the same block size in this schematic figure, the actual side-length of block is 25 km for the top 2 layers and 30 km for the lower layers

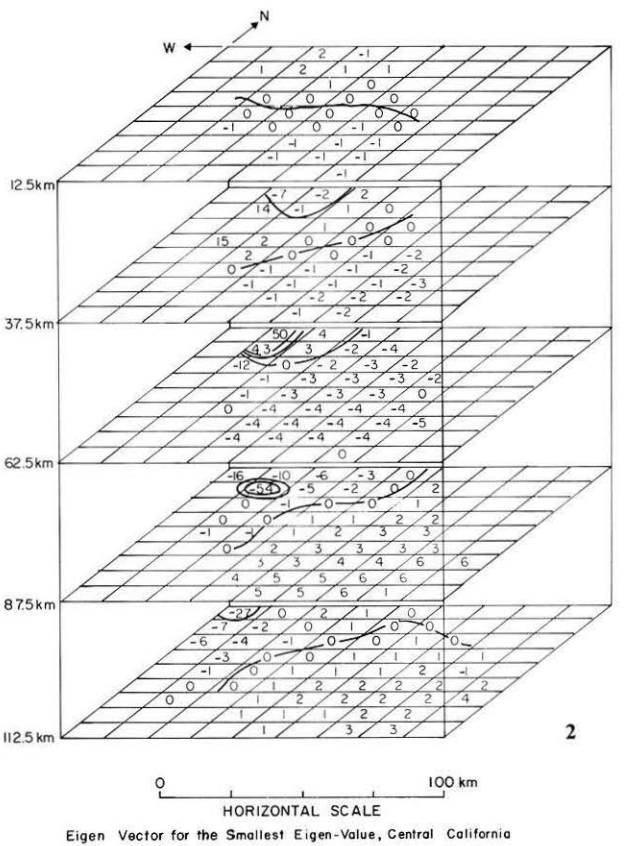

Fig. 2. Eigenvector for the smallest eigenvalue in the case of central California studied by Husebye et al. (1976). Otherwise, the caption of Figure 1 applies

top and second layers and 30 km for lower layers. The total number of blocks penetrated by at least one ray path is 205. The number of zero eigen-values is 12, of which 5 are due to the indeterminate constant for each layer and 7 due to coupling between two blocks in different layers, both effects mentioned earlier.

Figure 1 shows the eigen-vector with the largest eigen-value for the case of California described above. Each sheet represents a layer in the initial model, placed at its median depth. The number in each square represents component of eigen-vector for the block in %. They show very small values in the peripheral blocks. There are a peak and a trough in the middle of the array siting area which persist continuously from the top to the third layer and gradually die out. This is the pattern our data can determine most reliably and accurately.

In contrast to this pattern for the largest eigen-value, the components of eigen-vector for the smallest eigen-values are nearly zero everywhere except at one corner as shown in Figure 2. At that corner, they show large values with alternating signs between neighboring layers. This is the pattern which is most difficult to determine from our data. Small random errors can cause large fluctuation of this pattern.

There are 191 other eigen-vectors with spatial patterns of intermediate nature between the above two extreme ones. Together, they make up our inverse solution. The damped least squares method is designed to minimize the contribution from eigen-vectors with small eigen-values. It will smooth out the fluctuation due to random error at peripheral blocks, but lose resolution in defining vertical velocity variation because the eigen-vectors with large eigen-values tend to repeat a similar pattern for neighboring layers. Typically, we choose the damping factor θ^2 such that 2/3 of non-zero eigen-values are greater than θ. The resolution matrix V^TV corresponding to this choice of θ shows a diagonal element nearly uniform ($0.5 \sim 0.7$) throughout our model region except in the peripheral blocks where it drops below 0.5. The standard error of solution is typically $0.3 \sim 0.5\%$ so that the resultant slowness anomaly greater than 1% can be considered significant.

Summary of Results

Central California. A preliminary result based on about 1500 readings of teleseismic *P* residuals obtained from U.S. Geological Survey seismograph network revealed seismic velocity anomalies showing a remarkable correlation with the San Andreas fault (Husebye et al., 1976). They also showed that the lower limit of the root mean square of the true velocity perturbation is 3.1% for the crust-mantle under the array to a depth of 125 km. This estimate of velocity perturbation under the California array is close to the values for NORSAR (3.4%) and LASA (3.2%) array estimated by Aki et al. (1976a, b).

To test whether the large scale trend in velocity anomaly (parallel to the San Andreas fault) is due to the inhomogeneity inside or outside of our block model, Zandt and Aki (1976), using additional 2000 readings of *P*-time residuals compared solutions for two different assumptions on incident waves. The first solution was obtained using the incidence angle determined by the NOAA

epicenter and a standard earth (in this case, the Jefferys-Bullen travel time table). The second solution was obtained by assuming that the incident wave is the best-fitting plane wave locally determined from arrival times observed at the array. The small scale anomalies in two solutions are quite similar to each other. However, a large scale trend, such as the trend of San Andreas fault which exists in the first solution is absorbed into the adjusted direction of incident wave in the second solution and disappears.

Fortunately, we have here an independent set of data which allows us to choose between the two solutions. They are the P_n residuals from nuclear explosions in Nevada Test Site and near-by earthquakes studied by Kind (1972). His residuals were translated into a slowness anomaly map assuming a constant crustal thickness [to be compatible with our model] which showed a remarkable similarity to our first solution, obtained on the basis of a standard earth and NOAA epicenter.

The map of crustal velocity anomaly thus determined shows an excellent correlation with the geology map. Outcrops of granitic rocks on the west of San Andreas fault corresponds to high velocity anomalies. A low velocity wedge between Calaveras and San Andreas faults corresponds to sediments in the fault zone, the lowest velocity occurring where outcrops of serpentinite exist. Another prominent low velocity area near San Jose between the San Andreas and Calaveras faults also correlates well with an outcrop of serpentinite. Whatever the process which brought the serpentinite to surface might have weakened the crust causing the low velocity anomaly. Mt. Diablo, a site of strong gravity and magnetic anomaly is also marked by a high-velocity anomaly. The contrast between a peak and a trough of crustal velocity anomaly amounts to more than 10%.

On the other hand, anomalies in the mantle are weaker than in the crust. The trend of San Andreas fault appears to persist to a depth of about 55 km in the southern part of the array siting area and to deeper than 70 km in the north. At depths around 85 km, the pattern of anomalies become dominated by a low velocity anomaly with NE-SW trend, cutting across the San Andreas fault. One might speculate that the change of pattern occurred at the lithosphere-athenosphere interface. The depth of $55 \sim 85$ km is not unreasonable for the bottom of lithosphere in this area.

Yellowstone and Hawaii. Two other young and active areas are covered by seismic arrays operated by U.S. Geological Survey. The data from Yellowstone are analysed by Iyer (1976) and those from Hawaii by Ellsworth and Koyanagi (1976) using our method. Velocity anomalies in both areas correlate well with geology. In Yellowstone, a low-velocity column was found directly beneath the caldera, and it persists from the upper crust to a depth of 100 km. The diameter of low velocity column appears to increase with depth. The cause of this anomaly has been discussed by Eaton et al. (1975), who interprets it as a combination of a shallow silicic magma chamber underlain by a partially molten root. More recent data from extended array indicates persistence of low velocity anomaly to 200–250 km.

Kilauea, Hawaii makes an interesting contrast to Yellowstone. Here, the

crust immediately below the summit crater and rift zones (the site of fissure eruptions) is characterized by high velocity anomalies. This is consistent with previous refraction results as well as gravity anomalies, and may be explained by high density residuals accumulated along the path of magma ascent and/or by annealing effect of high temperature on cracks and pores of highly porous basaltic lava sheets which make up the island. The velocity anomalies in the mantle are again much weaker than those in the crust. Contrary to the high-velocity anomalies in the crust associated with eruption sites, an indication of low velocity anomaly trending NW-SE (parallel to the trend of Hawaiian Islands) was found in the mantle under volcanoes.

Now, let us leave young active areas where U.S. Geological Survey is interested in and go to arrays in old quiet areas chosen for monitoring remote underground nuclear testing. Interestingly, both in LASA (Montana, U.S.A.) and NORSAR (Norway), we fail to find a simple correlation between the crustal velocity anomaly and geology map.

Montana LASA. We used 3026 readings of *P*-time at 17 inner subarray centers for 178 teleseismic events tabulated by Chiburis and Ahner (1973). The details of analysis and result are give in Aki et al. (1976b). Within the LASA array, there are no conspicuous features in the gravity anomaly, sediment thickness or any other geological parameters. On the other hand, the seismic velocity anomaly is characterized by a strong trend in the N60°E direction, from the upper crust to a depth of 100 km. There is no indication of such anomalies in gravity or geology, except that a fault trace called "Weldon-Brockton fault zone" and the Bouguer gravity contour line run in the same N60°E direction in the north, immediately outisde the array area. The vertical cross-section of velocity anomalies along N30°W near the array center indicates a dipping low velocity sheet sandwiched between high velocity areas. Our speculation, here, is a large scale shear zone associated with building of the Rocky Mountains.

NORSAR. Finally, we come to NORSAR, where we used 2046 readings of *P*-times from 93 events recorded at 22 subarray centers. The data were tabulated by Berteussen (1974), and the result of analysis is described in Aki et al. (1976a). The eastern half of the array siting area is on that part of Baltic shield which was undisturbed either by the Caledonian orogeny or by Permian volcanism. This fact is reflected in the general pattern of higher velocity to east in the velocity anomaly map for all the layers. However, the most interesting geologic feature, Oslo graben (a site of Permian volcanism 200 million years ago) does not show up clearly in the velocity anomaly map of the crust. The trend of velocity anomaly in the crust appears to cut obliquely the boundary of Oslo graben, rather than running parallel to it.

In the mantle, we find a small-scale anomaly underlying the central part of the array, which becomes increasingly conspicuous with depth. In the bottom layer (depth range 96–126 km) we find a strong velocity variation amounting to 6% change in 50 km distance. The shape of this anomaly (low velocity) suggests a pipelike structure, and our speculation here is to associate this with the remains of magma ascent path through which the volcanic rocks of Oslo

graben were transported, somehow displaced from the location immediately beneath the graben area by lateral forces acting afterwards.

Conclusion

Although the above speculations are premature, some definite conclusions can be drawn from our results.

(1) Small-scale ($20 \sim 50$ km) strong velocity anomalies (amounting to more than 5% contrast) exist to a depth of at least 100 km.

(2) Velocity anomalies in the crust and geology map correlate well in young active areas such as California, Hawaii and Yellowstone. The correlation is poor in old, stable areas such as Montana and Norway.

(3) Deep-seated small-scale anomalies in the mantle are more conspicuous in old, stable areas.

Extending the 3-dimensional mapping of seismic velocity anomalies to the entire lithosphere of the earth may be useful for understanding the present and past tectonics. In particular, the shape and orientation of a velocity anomaly may be used to infer the stress history of the lithosphere.

Acknowledgements. Vigorous and persistent discussions with Anders Christoffersson and Eystein Husebye from the summer of 1974 through the spring of 1976 were the main source of material and ideas presented in this paper. Thanks are also due to Bill Ellsworth, H.M. Iyer, Bob Koyanagi, and George Zandt for making their results available to the author before publication. Bill Ellsworth read the manuscript and suggested improvements.

This work was supported by the U.S. Geological Survey under Contract 14-08-0001-G-339.

References

Aki, K., Christoffersson, A., Husebye, F.S.: Determination of the three-dimensional seismic structure of the lithosphere. J. Geophys. Res. in press, 1976a

Aki, K., Christoffersson, A., Husebye, E.S.: Three-dimensional seismic structure of the lithosphere under Montana LASA. Bull. Seism. Soc. Am. **66**, 501–524, 1976b

Berteussen, K.A.: NORSAR location calibrations and time delay corrections, NTNF/NORSAR. Sci. Rep. No. 2–73/74, 1974

Chiburis, E.F., Ahner, R.O.: LASA regional travel-time corrections and associated nodes. SDAC Technical Report No. 73–76, AD 774458 Teledyne, Geotech, Alexandria, Virginia 1973

Eaton, G.P., Christiansen, R.C., Iyer, M., Pitt, A.M., Mabey, D.R., Blank, H., Zietz, I., Gettings, M.E.: Magma beneath Yellowstone National Park. Science **188**, No. 4190, 787–796, 1975

Ellsworth, W.L., Koyanagi, R.Y.: Three-dimensional crust and mantle structure of Kilauea volcano, Hawaii. Submitted to J. Geophys. Res. 1976

Franklin, J.N.: Well-posed stochastic extension on ill-posed linear problems. J. Math. Anal. Appl. **31**, 682–716, 1970

Husebye, E.S., Christoffersson, A., Aki, K., Powell, C.: Preliminary results on the 3-dimensional seismic structure of the lithosphere under the USGS Central California Seismic Array. Geophys. J.R. Astr. Soc. **46**, 319–340, 1976

Iyer, H.M.: Personal communication. 1976

Kind, R.: Residuals and velocities of P_n waves recorded by the San Andreas Seismograph network. Bull. Seism. Soc. Am. **62**, 85–100, 1972

Lancsoz, C.: Linear differential operators, pp. 564. London-New York: Van Nostrand 1961

Zandt, G., Aki, K.: Lateral velocity anomalies associated with the San Andreas fault, central California. Trans. Am. Geophys. Union, EOS **57**, 283, 1976

Received October 12, 1976; Revised Version December 6, 1976