INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

H4.SMR/942-13

Third Workshop on 3D Modelling of Seismic Waves Generation Propagation and their Inversion

4 - 15 November 1996

Surface Wave Tomography: Theory and Applications

T. B. Yanovskaya
St. Petersburg State University
Institute of Physics
St. Petersburg
Russian Federation

SURFACE WAVE TOMOGRAPHY: THEORY AND APPLICATIONS

T.B. Yanovskaya

Contents

- Basic principles of surface wave tomography
- Resolving power
- Tomography based jointly on the data on phase and group velocities
- Estimation of lateral velocity variations and anisotropy
- Azimuthal tomography

1. Basic principles of surface wave tomography

Seismic tomography is reconstruction of 3D Earth structure from the data obtained along different wave paths. Usually determined are seismic wave velocities, and the data are travel times, though some other characteristics of the waves, even the wave forms, may be used for the purpose. Surface wave tomography is based on the data on surface wave observations.

Surface waves propagate in upper layers of the Earth, and the thickness of the layer, in which the most part of their energy is concentrated, depends on the wavelength: the larger is the wavelength, the deeper the wave is penetrated. Therefore characteristics of surface waves for different wavelengths, or periods, contain information on 3D structure of the upper part of the Earth - the Earth's crust and upper mantle.

Thus, surface waves may be used for the abovementioned purpose - reconstruction of a 3D distribution of elastic parameters in the upper part of the Earth. However, the peculiarities of surface wave propagation lead to the corresponding peculiarities of the tomography methods.

It is clear that for 3D tomography we must use the data for a sufficiently wide range of periods. It can be done by two approaches. One is the use of the data for the whole range of periods altogether, and to reconstruct the 3D structure directly. Another one is to divide the tomographic reconstruction into two steps: at the first step the data for a fixed period are used, and 2D tomography problem is solved to determine a 2D distribution of surface wave characteristics - phase or group velocities. Then, when this problem has been solved for different periods, vertical velocity sections in different points of the area can be obtained by solving 1D inverse problem. In other words, the initial 3D problem is divided into two problems: properly 2D tomography problem for phase or group surface wave velocities, and 1D inverse problem for dispersion curves.

Example of the first approach is the so-called 'waveform tomography' proposed by Nolet (1990). This method is based on determination of a 3D model, for which the calculated wave forms fit to those observed at different stations from different sources. Adjustment of the wave forms is practically equivalent to the adjustment of their phases, i.e. of the dispersion curves of group velocities. But this approach is too difficult for the practical use. It requires to calculate the waveforms in 3D inhomogeneous media, and to solve the inverse problem with too many parameters.

The second approach is easier for practice, therefore it is widely used in seismological studies. Most of studies are restricted with the first step - solving the 2D problem for determination of lateral variations of phase or group velocities, because they are sufficient for inferences about lateral variations of the lithosphere structure. It is obvious that 2D tomography problem is much easier than 3D. Therefore the development of this approach is connected with elaboration of the methods specific for 2D tomography.

2. Methods for solving 2D tomography problem

2.1. General approach

The data obtained from surface wave observations, which can be used for tomography, are:

group velocities,
phase velocities,
azimuthal anomalies.

The data on phase or group velocities are the same as group or phase travel times, and after linearizarion these data are reduced to time delays. The time delays δt_i are related to the 2D phase or group velocity pattern $V(\mathbf{r})$ ($\mathbf{r}=(x,y)$ or (φ,λ)) by the following functional

$$\delta t_i = t_i - t_{0i} = \int_{L_i} \frac{ds}{V(\mathbf{r})} - \int_{L_{0i}} \frac{ds}{V_0(\mathbf{r})} \cong \int_{L_{0i}} \delta\left(\frac{1}{V}\right) ds \qquad (1)$$

where $V_0(\mathbf{r})$ is velocity in the starting model.

The equation (1) may be rewritten in the standard form

$$d_i = \int_S G_i(\mathbf{r}) m(\mathbf{r}) d\mathbf{r} \qquad (2)$$

where $m(\mathbf{r})$ is the unknown model - relative perturbation of slowness $\delta V^{-1}(\mathbf{r}) / V_0^{-1}(\mathbf{r})$, and $G_i(\mathbf{r})$ is the data kernel, which is singular on the i-th ray, equal to zero elsewhere, and satisfying the constraint

$$\int_S G_i(\mathbf{r}) d\mathbf{r} = \int_{L_{0i}} \frac{ds}{V_o(\mathbf{r})} = t_{oi} \qquad (3)$$

The relationship between azimuthal anomaly and the model may be written also in the form (2), but the data kernel $G_i(\mathbf{r})$ will be different (see below).

If the data are exact, the problem is to determine $m(\mathbf{r})$ from a finite set of equations (2). If they have errors, $m(\mathbf{r})$ should be determined by minimization of the functional

$$\sum_{ij} \left(d_i - \int_S G_i(\mathbf{r}) m(\mathbf{r}) d\mathbf{r} \right) R_{ij}^{-1} \left(d_j - \int_S G_j(\mathbf{r}) m(\mathbf{r}) d\mathbf{r} \right) \qquad (4)$$

where **R** is covariance matrix of the data errors.

Since a function cannot be determined from a finite data set, some a priori constrains should be imposed to the unknown function. The constrains may be of different kind, but finally the solution may be written in the following general form

$$m(\mathbf{r}) = \sum_{j=1}^{M} \lambda_j \psi_j(\mathbf{r}) \qquad (5)$$

where $\psi_j(\mathbf{r})$ are some basis functions. The basis functions may be either assumed a priori, or constructed proceeding from the given data set. Once the basis functions are assumed, the coefficients λ_j may be determined from a linear system of equations, which is obtained by substitution of (5) into (2) or (4).

2.2 Alternatives for choice of the basis functions

The basis functions $\psi_j(\mathbf{r})$ may be either chosen *a'priori*, or constructed from some properties assumed about the model $m(\mathbf{r})$.

1. A *'priori* choice of the basis functions. In surface wave tomography usually two approaches are applied. If the wave paths cover the whole globe, the model $m(\varphi,\lambda)$ is represented as a series in spherical harmonics, so that the basis functions are just spherical harmonics. The number of the basis functions is chosen to be less than the number of the data, then the coefficients by the functions are determined from minimization of the functional (4). However, if the paths cover the globe non-uniformly - and this situation is usual, - then the solution will have artificial details in the parts poorly covered by the paths, and on the contrary, some details in the regions crossed by a large number of paths will be smoothed.

Another approach is used in regional studies - it is so-called *regionalization:* the area under investigation is divided into some regions R_j (j=1,2,..,M), which are assumed to be laterally homogeneous on the basis of geomorphological, geotectonical or some other geophysical data. The velocities within each region are assumed to be constants (for a given period), so that the problem is reduced to estimation of a number of parameters, which are the unknown slowness corrections

. The basis functions in this case are defined as follows:

$$\psi_j(\mathbf{r}) = \begin{cases} 1 & \text{if } \mathbf{r}\subset R_j \\ 0 & \text{otherwise} \end{cases} \qquad (6)$$

A drawback of the latter approach is that it requires *a'priori* regionalization, which is usually made on the basis of differences in surface features, though they may not reflect the lateral heterogeneity at depth.

2. *Determination of the basis functions from a'priori assumptions about the model.* One method was proposed by Tarantola & Nersessian (1984) on the basis of Bayesian approach. If the *a'priori* covariance function of the model $C(\mathbf{r},\mathbf{r}')$ is assumed, $m(\mathbf{r})$ is determined by minimizing the functional:

$$\int_S \int_S m(\mathbf{r}) C^{-1}(\mathbf{r},\mathbf{r}') m(\mathbf{r}')\,d\mathbf{r}\,d\mathbf{r}' + (\mathbf{d}-\mathbf{Gm})^{\mathrm{T}} \mathbf{R}_d^{-1} (\mathbf{d}-\mathbf{Gm}) \qquad (7)$$

where $\mathbf{Gm}$ is a vector with components $\int_S G_i(\mathbf{r}) m(\mathbf{r})\,d\mathbf{r}$.

The solution is $\mathbf{m} = \mathbf{C}\mathbf{G}^T(\mathbf{R}_d + \mathbf{G}\mathbf{C}\mathbf{G}^T)^{-1}\mathbf{d}$, where the following notation is implied:

$$(\mathbf{G}\mathbf{C}\mathbf{G}^T)_{ij} = \iint G_i(\mathbf{r})C(\mathbf{r},\mathbf{r}')G_j(\mathbf{r}')d\mathbf{r}\,d\mathbf{r}'$$

The basis functions are

$$\psi_j(\mathbf{r}) = \int_S C(\mathbf{r},\mathbf{r}')G_j(\mathbf{r}')d\mathbf{r}' \qquad (8)$$

The *a'priori* covariance function of the model is assumed usually as Gaussian

$$C(\mathbf{r},\mathbf{r}') = \sigma^2 \exp\left(-\frac{|\mathbf{r}-\mathbf{r}'|^2}{2L^2}\right)$$

where L is a correlation length.

It should be noted that this method is valid both for 2D and 3D tomography problems.

Another method proposed by Ditmar & Yanovskaya(1987, 1990) is valid only for 2D tomography problems. It is based on the assumption of smoothness of the solution according to the criterion

$$\int |\nabla m|^2 d\mathbf{r} = \min \qquad (9a)$$

$$m = \text{const} \quad \text{at} \quad |\mathbf{r}| \to \infty \qquad (9b)$$

The condition (9a) provides continuity of $m(\mathbf{r})$, though the derivatives of this function are discontinuous at the rays. For accurate data the condition analogous to (9b) is

$$(\mathbf{d}-\mathbf{G}\mathbf{m})^T\mathbf{R}_d^{-1}(\mathbf{d}-\mathbf{G}\mathbf{m}) + \alpha \int |\nabla m|^2 d\mathbf{r} = \min, \qquad (10)$$

where α is parameter of regularization. The basis function in Cartesian coordinates satisfying (9a) or (10) are following:

$$\psi_i(\mathbf{r}) = \begin{cases} \iint \ln|\mathbf{r}-\mathbf{r}'| G_i(\mathbf{r}')d\mathbf{r}' & i = 1,2,..\ N \\ 1 & i = N+1 \end{cases} \qquad (11)$$

In case of accurate data the coefficients by the basis functions are determined from the following system of equations:

$\mathbf{S}\lambda + c\mathbf{t}_0 = \mathbf{d}$

$\lambda^T\mathbf{t}_0 = 0$

where $\lambda^T = (\lambda_1, \lambda_2, \ldots \lambda_N,)$ are the coefficients by the N first functions, and c is the coefficient by the $(N+1)$-th function, the matrix $\mathbf{S}$ being defined as

$$S_{ij} = \int_{L_i}\int_{L_j} \ln|\mathbf{r}_i - \mathbf{r}_j| \frac{ds_i}{V_0(\mathbf{r}_i)} \frac{ds_i}{V_0(\mathbf{r}_i)} \qquad (12)$$

If the data are inaccurate, the coefficients are determined from the system

$(\mathbf{S}+\alpha\mathbf{R}_d)\lambda + c\mathbf{t}_0 = \mathbf{d}$

$\lambda^T\mathbf{t}_0 = 0$

Examples of the basis functions (8) and (11) are shown in Figs. 1,2.

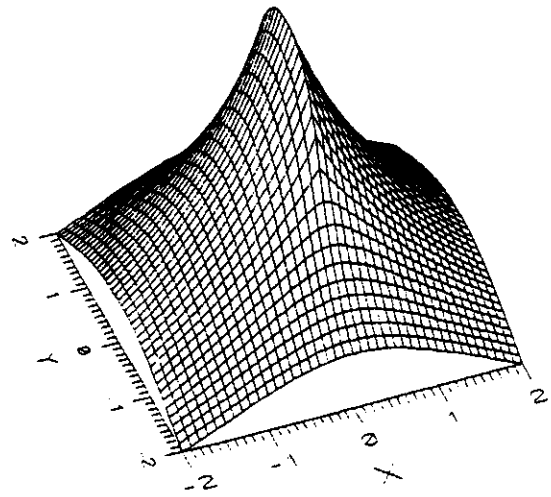

Figure 1. Example of the basis function $\psi(x, y)$ for a ray along the y-axis between the points $y = -1$ and $y = +1$. It is clear that the derivative $d\psi/dx$ is discontinuous on the ray.

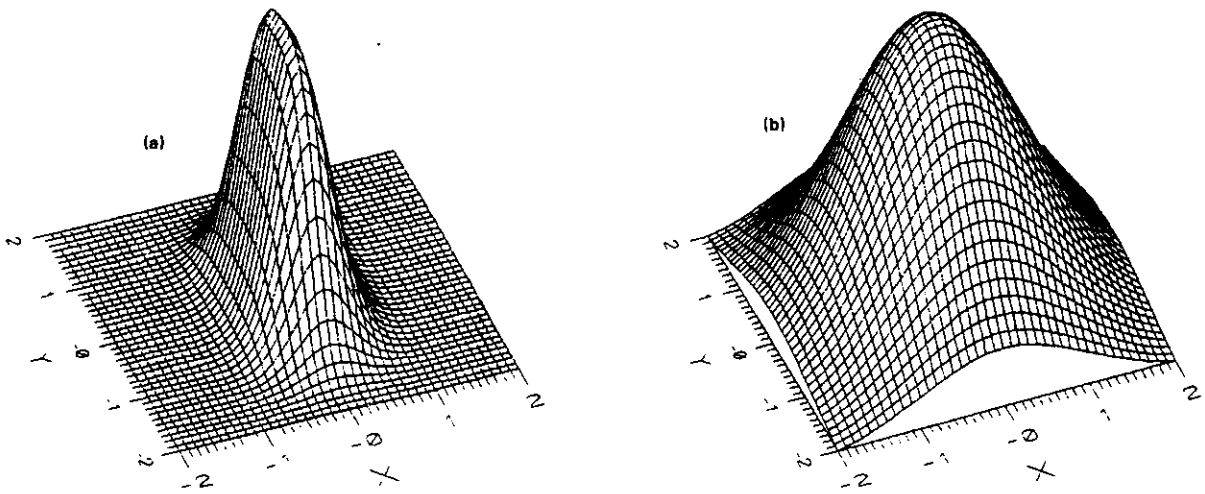

Figure 2. Examples of the basis functions (16) for the same ray as in Fig. 1 for two different values of L: (a), $L = 0.333$; (b), $L = 1.0$.

3. Resolving power of the data

A solution of the tomography problem is not unique, because the data (functionals along the rays) do not contain information on the velocities in all points, so that the solution is always averaged over some area. Therefore estimation of resolving power of the data is an important problem in seismic wave tomography. A knowledge of the resolution allows to estimate sizes of anomalies , which can be revealed reliably from a given set of the data. Also it allows to conclude if the anomalies obtained in the solution are realistic or artificial.

In general the resolving power of the data may be estimated by the *averaging kernel*. A solution of the tomography problem is sought in a general form

$$m(\mathbf{r})=\lambda^{T}\psi(\mathbf{r})$$

where $\psi^{T}(\mathbf{r})=\{\psi_1(\mathbf{r}), \psi_2(\mathbf{r}),\ldots. \psi_M(\mathbf{r})\}$ - basis functions. Since the problem is reduced to a linear system of equations in respect to the coefficients λ_j, in the right-hand side of which are the data d_i, or their linear combination, the solution for $m(\mathbf{r})$ may be also expressed as a linear combination of the data:

$$\hat{m}(\mathbf{r}) = \sum a_i(\mathbf{r}) d_i \qquad (13)$$

This is one of the solutions satisfying the data

$$d_i = \int G_i(\mathbf{r}) m(\mathbf{r}) d\mathbf{r} \qquad (14)$$

Therefore the solution may be written as

$$\hat{m}(\mathbf{r}) = \sum_i a_i(\mathbf{r}) \int G_i(\mathbf{r}') m(\mathbf{r}') d\mathbf{r}', \qquad (15)$$

where $m(\mathbf{r}')$ is any solution of (14).

Otherwise (15) may be written as $\hat{m}(\mathbf{r}) = \int A(\mathbf{r},\mathbf{r}') m(\mathbf{r}') d\mathbf{r}'$, where $A(\mathbf{r},\mathbf{r}')$ is the averaging kernel, which is expressed as a linear combination of data kernels.

In tomography problems the data kernels are singular at the rays, consequently the averaging kernel is also singular. Therefore it is inconvenient to characterise the resolution in each point by the averaging kernel directly: besides that it is singular, it should be determined in each point r, so that it requires a tremendous amount of computations.

In some studies the resolution is estimated by calculation of the solution for some synthetic models (Snieder, 1988). Usually the model contains a set of low and high velocity anomalies of a certain size alternating in chess-board order. The time delays are calculated for the same paths as in the real data set. If these anomalies are appeared in the solution, this means that in the real problem the anomalies of the same size would be resolved.

The advantage of such approach is that it is unnecessary to obtain the inverse operator in an explicit form: the iterative methods may be used for solving the problem. A main drawback is that it is difficult to conclude, anomalies of what size can be resolved in different parts of the area.

For estimation of the resolution in the case, when the solution is determined from the criterion of smoothness, a method analogous to that proposed by Backus & Gilbert can be applied (Ditmar & Yanovskaya, 1987). The resolution in each point is estimated by a linear size of the averaging area. It has been shown that the solution minimizing (9a) is equivalent to that obtained from a criterion of proximity of the averaging kernel to the 2D delta-function (so-called δ-ness criterion). For singular kernel the δ-ness criterion

may be replaced by a criterion of proximity of the integral of the kernel to the integral of the delta-function (in 1D case to the Heaviside function). In 2D case this criterion is following:

$$s^2(\mathbf{r}) = \int \left| E(\mathbf{r},\mathbf{r}') - H(\mathbf{r},\mathbf{r}') \right|^2 d\mathbf{r}' = \min \qquad (16)$$

under the normalization condition

$$\int A(\mathbf{r},\mathbf{r}')\,d\mathbf{r}' = 1,$$

where $div E(\mathbf{r},\mathbf{r}') = A(\mathbf{r},\mathbf{r}')$, $div H(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r}'-\mathbf{r})$.

To evaluate a size of the averaging area, we consider the kernel which would be a constant equal to $1/\pi R^2$ within a circle of radius R centred at r and vanish outside the circle. It is easy to calculate $s^2(\mathbf{r})$ for such kernel, which is obviously a function of R. Then the value of R, which provides the same $s^2(\mathbf{r})$, may be accepted as an estimate of a linear size of the averaging area. It may be shown that

$$R = \exp(3/4 - \mathbf{a}^T S \mathbf{a} + 2\mathbf{a}^T \psi) \qquad (17)$$

where a is defined by (13), $\psi(\mathbf{r})$ is the vector formed by the N basis functions (11), and the matrix S is defined by (12).

Fig 4 shows the distribution of $R(x,y)$ for the pattern of paths shown in fig.3. The effective radius of the averaging area is the smallest in those parts of the region, which are covered most densely by the rays.

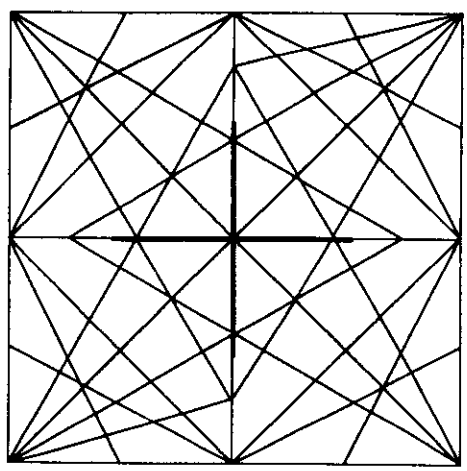

Fig.3

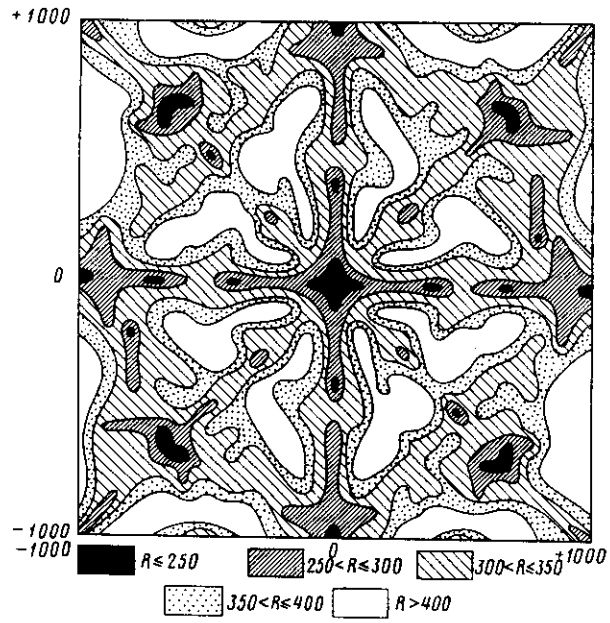

Fig.4

A drawback of the estimation of the resolution by the effective radius of the averaging area is that it does not allow to conclude about the resolving power in different directions. Indeed, when the rays are oriented predominantly in one direction, the resolving power is low in this direction and high in the perpendicular direction.

Below we consider an approach for estimation of the resolution depending on azimuth.

Let us consider a plane area Σ $(0<x<X,\ 0<y,Y)$ covered by the rays. The data are travel time residuals δt_i. As mentioned above, a solution of the tomography problem in a point X_0, Y_0 may be represented as a linear combination of the data:

$$m(X_0, Y_0) = \sum_i a_i(X_0, Y_0)\delta t_i$$

The averaging kernel in the point X_0, Y_0 is expressed as follows:

$$A(X, Y; X_0, Y_0) = \sum_i a_i(X_0, Y_0)G_i(X, Y) \qquad (18)$$

where the data kernels $G_i(X, Y)$ satisfy (3).

Now we determine a criterion for estimating the resolution in a point X_0, Y_0 *along* a fixed direction. Without a loss of generality this direction may be assumed coinciding with x-axis direction.

Integrating (18) with respect to y-coordinate over the interval $(0, Y)$, corresponding to the area covered by the rays, we obtain a function of x, depending on the coordinates X_0, Y_0 :

$$F(X; X_0, Y_0) = \sum_i a_i(X_0, Y_0)\int_0^Y G_i(X, Y)dy \qquad (19)$$

This function characterises a rate of smoothing of the solution along the x-direction. Therefore a resolution in this direction may be determined by a deviation of this function from the delta-function $\delta(X - X_0)$. But since they both are singular, it is reasonable to replace this deviation by a deviation of the integral of this function

$$\Phi(X; X_0, Y_0) = \int_0^X F(X'; X_0, Y_0)dX' \quad (20)$$

from the Heaviside function $H(X - X_0)$. Substituting (19) into (20) we obtain:

$$\Phi(X; X_0, Y_0) = \sum_i a_i(X_0, Y_0)\tau_i(X) \ , \qquad (21)$$

where $\tau_i(X)$ is travel time along the i-th ray up to the point X. Example of such function for the straight-forward ray with the end-points X_i^1 , X_i^2 is shown in fig.5.

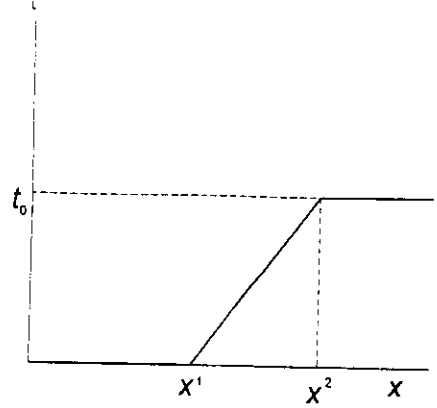

Fig.5

It is reasonable to determine the deviation of the function (21) from the Heaviside function by the mean square residual

$$s(X_0, y_0) = \int_0^X \left(\Phi(X; X_o, y_0) - \varepsilon(X - X_0)\right)^2 dX \qquad (22)$$

The less s, the less difference between $\Phi(X; X_0, y_0)$ and the Heaviside function, and the higher resolution in X -direction. The value of s may be used for estimating the effective averaging length analogously to that as in Backus-Gilbert method for 1D inverse problems.

Let the averaging kernel differs from zero and equal to a constant value within the rectangle

$X_0 - \frac{a}{2} < X < X_0 - \frac{a}{2}$, $y_0 - \frac{b}{2} < y < y_0 - \frac{b}{2}$. Assuming the kernel to be normalized, this constant should be equal to $\frac{1}{ab}$. It is easy to show that $s(x_0, y_0)$ for such kernel turns out to be equal to $\frac{a}{12}$. Consequently, $12s$ may be taken as an estimate of the width of the averaging area in X- direction.

As mentioned above, the direction of x-axis may be chosen arbitrarily. Therefore the width of the averaging area can be estimated in any direction by rotating the X-axis, i.e. a may be determined as a function of azimuth ϑ .

Estimation of the resolving power by the effective radius of the averaging area (Ditmar & Yanovskaya, 1987) is equivalent to the assumption that a linear size of the area is equal in all directions, therefore it can be estimated by averaging $a(\vartheta) / 2$ over the whole interval of ϑ . The next step in detalization of the resolving power is approximation of $a(\vartheta)$ by a function $a + b\cos(2\vartheta - 2\phi)$. Then in the azimuths $\vartheta = \phi$ and $\vartheta = \phi + \pi$ the resolution is the worst (the linear size L of the averaging area in these azimuths is $a + b$), and in azimuths $\vartheta = \phi + \pi / 2$ and $\vartheta = \phi + 3\pi / 2$ it is the best ($L = a - b$). If $b \approx 0$, the resolution is approximately one and the same in all azimuths, and the estimation of the resolving power by the radius of the averaging area is justified. Large values of $/ a$ indicate to predominance of the rays of a certain direction, and consequently to insufficiency of the data for tomographic reconstruction.

Thus, in this approach the effective averaging area is of elliptic shape, with axes $a+b$ and $a-b$, maximum axis being in the azimuth ϕ.

Fig.6 shows effective averaging areas for two patterns of rays. In Fig 6a the areas are stretched in the predominant directions of the rays. In fig.6b, where the rays are oriented approximately uniformly, the areas are close to circles.

Figs.7a,b show distributions of the value $2b / a$ for the two sets of rays. These values are much smaller in the latter case.

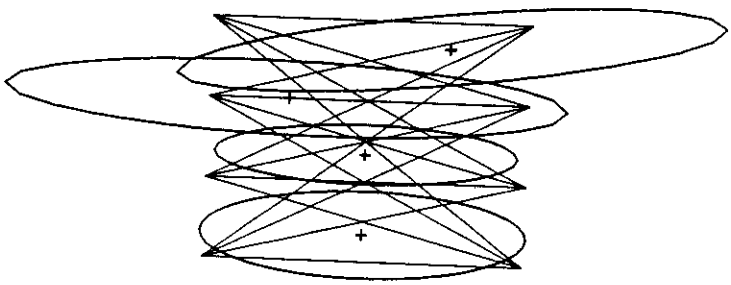

Fig.6a

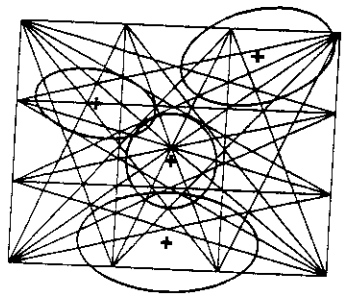

Fig.6b

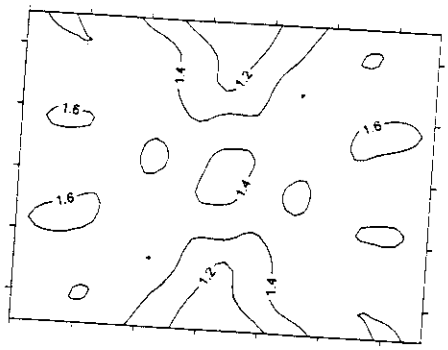

Fig.7a

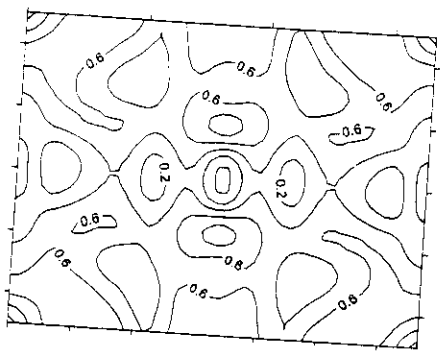

Fig.7b

4. Tomography based jointly on the data on phase and group velocities

The method described in the last section assumes knowledge of either phase or group velocities of surface waves along different paths for a fixed period. The solution will represent the velocity distribution as a function of the lateral coordinates for the period. This approach is not however always practicable for the following reasons.

First, observations do not always provide enough data for the particular period under consideration, while data may be available for near periods along a fairly large number of paths, and they contain certain information on the velocity for that period which the above approach ignores.

Secondly, if we remember that the resulting velocity distributions $V(x, y)$ ($V = C$ or U) are to be used for determining crustal and upper mantle structure, then it is fairly obvious that the method should be applied to find phase velocity distributions, since the use of group velocity dispersion in this problem may lead to much greater nonuniqueness associated with the fact that a dispersion curve of group velocity may have infinitely many curves of phase velocity corresponding to it, hence infinitely many velocity structures. At the same time, group velocities are much simpler to determine from observations than phase velocities; there are quite large amounts of data of this kind accumulated for different areas. This accounts for the fact that it is group velocities which are used in the many papers dealing with lateral variations of surface wave velocities (Yanovskaya, 1982; Yanovskaya and Nikolova, 1984; Sabitova and Yanovskaya, 1986; Dmitrieva *et al.*, 1986), even though the conclusions they contain on lateral inhomogeneities cannot be considered sufficiently substantiated.

Determination of phase rather than group velocity distributions is important, also because this enables one to incorporate nonlinearity of inverse problem in finding $V(x, y)$ from the set of functionals (6.30). We recall that (6.30) can be reduced to the linear functional (6.32) in which the integration is done along a straight line, provided the ray shape can be assumed to be determined by the starting model, which is $V_0 =$ constant in our case. However, if the lateral variation of velocity is not small enough, one should accomodate ray curvature; this can be done by successive approximations, the initial approximation at each iteration being the solution obtained at the last iteration, i.e., the path of integration in (6.32) is taken to correspond to the velocity distribution obtained at the last iteration. Now since ray shape is controlled by the distribution of phase velocity, then obviously, successive approximations cannot be constructed, unless it is the phase velocity that one obtains at the last iteration. However, these drawbacks of the above approach can be obviated, if one recalls the well-known relation between phase and group velocity

$$U^{-1}(\omega) = C^{-1}(\omega) + \omega \frac{\mathrm{d}C^{-1}(\omega)}{\mathrm{d}\omega} \tag{6.81}$$

so that the dispersion curve of phase velocity can be completely determined from that of group velocity, provided the constant of integration is known. In the case of a single dimension it is sufficient to know phase velocity at a single value of

frequency. This fact can form the basis of a technique for determining phase velocity distributions for different frequencies in laterally varying structures from the values of group velocity along different paths, provided we know a few values of phase velocity along some paths in the study area. The technique in question was proposed by Yanovskaya *et al.* (1988).

The raw data are phase $C_i^{(k)}$ $(i = 1, 2, \ldots, m_k)$ and/or group velocities $U_j^{(k)}$ $(j = 1, 2, \ldots, n_k)$ for periods $T_i^{(k)}$) along paths $L_k (k = 1, 2, \ldots, K)$ traversing the study area. A path is as before defined by giving the coordinates of the initial and end point; the problem will be solved in the linearized formulation, i.e., paths are assumed to be straight lines L_{0k}. The data can be converted into travel time residuals to be related to the unknown distributions of phase and group velocities $C(x, y, T)$ and $U(x, y, T)$ through

$$\begin{aligned} \int_{L_{0k}} \delta C^{-1}(x, y, T_i^{(k)})\, \mathrm{d}l &= \delta t_{ik}^{ph} \\ \int_{L_{0k}} \delta U^{-1}(x, y, T_j^{(k)})\, \mathrm{d}l &= \delta t_{jk}^{gr} \end{aligned} \tag{6.82}$$

where $\delta C^{-1} = C^{-1}(x, y, T) - C_0^{-1}(T)$, $\delta U^{-1} = U^{-1}(x, y, T) - U_0^{-1}(T)$, $C_0(T)$ and $U_0(T)$ being dispersion curves of phase and group velocity, respectively, averaged over the area. It should also be remembered that C and U are related by (6.81).

The idea of the method is that the phase slowness as a function of frequency $\omega = 2\pi/T$ is fitted by a polynomial whose coefficients are unknown functions of the lateral coordinates

$$C^{-1}(x, y, \omega) = \sum_{q=0}^{N} a_q(x, y)\, \omega^q. \tag{6.83}$$

Accordingly, group slowness is also fitted by a polynomial of the same degree:

$$U^{-1}(x, y, \omega) = \sum_{q=0}^{N} (1 + q)\, a_q(x, y)\, \omega^q. \tag{6.84}$$

To find $C(x, y, \omega)$, we must determine the coefficients a_q which are now functions of the spatial coordinates x, y only. This problem can in turn be reduced to that discussed in the last section.

We represent the 'mean' dispersion curve $C_0(\omega)$ in the form (6.83) again:

$$C_0^{-1}(\omega) = \sum_{q=0}^{N} a_{0q}\, \omega^q \tag{6.85}$$

where the a_{0q} are constants to be determined from $C_0(\omega)$. Accordingly, $U_0(\omega)$ is also expressible in terms of the same corefficients:

$$U_0^{-1}(\omega) = \sum_{q=0}^{N} a_{0q}(1 + q)\, \omega^q. \tag{6.86}$$

Equations (6.82) for the kth path can now be written as

$$\sum_{q=0}^{N} (\omega_i^{(k)})^q \, a_{0q} \int_{L_{ik}} m_q(x, y) \, dl = \delta t_{ik}^{ph} (i = 1, 2, \ldots, m_k)$$

$$\sum_{q=0}^{N} (\omega_j^{(k)})^q \, a_{0q}(1 + q) \int_{L_{jk}} m_q(x, y) \, dl = \delta t_{jk}^{gr} (j = 1, 2, \ldots, n_k) \tag{6.87}$$

where $m_q(x, y) = (a_q(x, y) - a_{0q})/a_{0q}$. These equations for all paths $k = 1, 2, \ldots, K$ can be combined to form the set of equations

$$\sum_{q=0}^{N} b_{rq} \int_{L_{rk}} m_q(x, y) \, dl = \gamma_r, \qquad r = 1, 2, \ldots, M \tag{6.88}$$

where $\gamma_r = \delta t_{ik}^{ph}$ or δt_{jk}^{gr}, $k = k(r)$, $b_{rq} = a_{0q}(\omega_{i(j)}^{(k)})^q(1 + pq)$, $p = 0(1)$ corresponding to phase (group) velocity; $M = \Sigma_k(m_k + n_k)$.

One can find $m_q(x, y)$ from (6.88) using a method which is an extension of that described in the last section.

We assume the $m_q(x, y)$ to be components of a vector fuction $\mathbf{m}(x, y)$ and demand this function to obey a smoothness requirement similar to (6.35):

$$\int_S \int \|\nabla \mathbf{m}\|^2 \, dx \, dy = \min \tag{6.89}$$

where $\nabla \mathbf{m}$ is a matrix of the derivatives $\partial m_q/\partial x_k$. We also assume $\mathbf{m}(x, y)$ to be bounded at infinity. It then turns out that each of the $m_q(x, y)$ must satisfy Poisson's equation (6.37); its solution has the form

$$m_q(\mathbf{r}) = \sum_{i=1}^{M} \lambda_i \, b_{iq} \int_{L_{ik}} \ln |\mathbf{r} - \mathbf{r}_i| \, dl_i + C_q. \tag{6.90}$$

We have taken path number to correspond to equation number in (6.88), so that different indices may correspond to one and the same path.

The boundedness of $\mathbf{m}(\mathbf{r})$ at infinity yields the requirement that the λ_i must satisfy

$$\sum_{i=1}^{M} \lambda_i \, b_{iq} \, l_i = 0 \tag{6.91}$$

where l_i is the length of the ith path.

We define matrices $\mathbf{S}$ and $\mathbf{D}$ as follows:

$$S_{ij} = \sum_{q=0}^{M} b_{iq} \, b_{jq} \int_{L_{ik}} \int_{L_{jl}} \ln |\mathbf{r}_i - \mathbf{r}_j| \, dl_i \, dl_j, \qquad 1 \leqslant i, j \leqslant M \tag{6.92}$$

$$D_{ij} = b_{ij} \, l_i, \quad 1 \leqslant i \leqslant M, \quad 0 \leqslant j \leqslant N. \tag{6.93}$$

The coefficients λ_i (the vector Λ) and the constant terms C_q (the vector $\mathbf{C}$) satisfy

the following set of equations:

$$\mathbf{S}\Lambda + \mathbf{D}\mathbf{C} = \boldsymbol{\gamma} \qquad \mathbf{D}^T\Lambda = 0. \tag{6.94}$$

The set (6.94) is similar to (6.43), (6.45). If we now define a matrix $\mathbf{K}(\mathbf{r})$:

$$K_{iq}(\mathbf{r}) = b_{iq} \int_{L_m} \ln|\mathbf{r} - \mathbf{r}_i| \, dl \qquad 1 \leqslant i \leqslant M, \quad 0 \leqslant q \leqslant N \tag{6.95}$$

we obtain a solution for $\mathbf{m}(\mathbf{r})$ that is an extension of (6.48):

$$\mathbf{m}(\mathbf{r}) = \mathbf{K}^T\mathbf{S}^{-1}\boldsymbol{\gamma} + (\mathbf{I} - \mathbf{K}^T\mathbf{S}^{-1}\mathbf{D})(\mathbf{D}^T\mathbf{S}^{-1}\mathbf{D})^{-1}\mathbf{D}^T\mathbf{S}^{-1}\boldsymbol{\gamma}. \tag{6.96}$$

Note that $\mathbf{S}$ is degenerate, if there are more data points for a path than the coefficients of the fitting polynomial. This means that equations (6.88) must be satified approximately rather than exactly. If we assume them to be satisfied in the sense of the minimum squared residuals, the appropriate solution is (6.96) again, but the inverse matrix $\mathbf{S}^{-1}$ is now to be the generalized inverse $\mathbf{S}^+$ based on the singular value decomposition of $\mathbf{S}$.

Model example. This method has been tested on the following model example. The distribution of phase velocity in the region $-1000 \leqslant x \leqslant 1000$, $-1000 \leqslant y \leqslant 1000$ is given as

$$C(x, y, \omega) = 1.5(3.0 - \tan^{-1}(2\omega^{3/2} H(x, y)) \tag{6.97}$$

where $H(x, y) = 1 + \exp[-(x/1000)^2 - 3(y/1000)^2]$. Phase and group travel times were calculated along the 28 paths shown in Figure 6.9 for $0.2 \leqslant \omega \leqslant 0.5$, and a sample of 100 travel times was used to determine phase velocity in the study area.

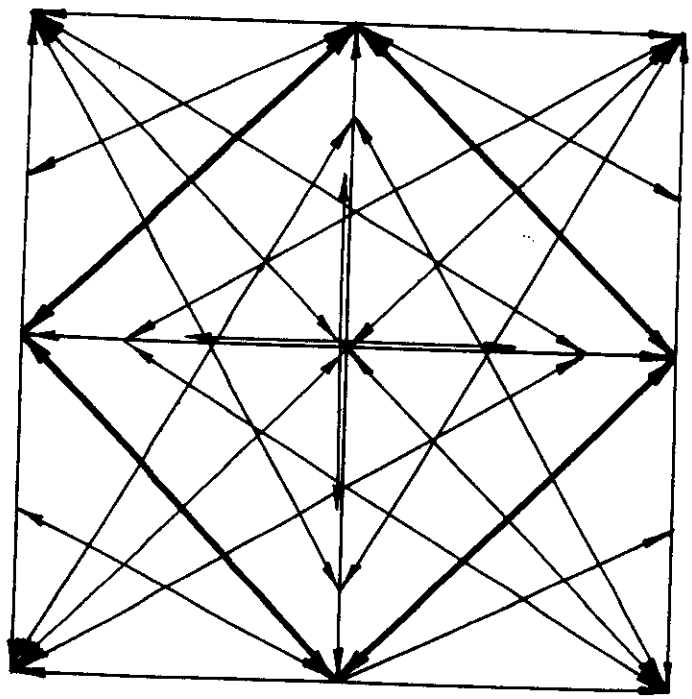

Fig. 6.9. A pattern of paths the travel times for which have been used in a test inversion. Solid lines are paths for which phase velocities were known.

The number of data points for a path was 3 to 5, the phase slowness was fitted by a straight line, hence $N = 1$. The phase travel times were taken only for the 4 paths indicated by heavy lines in Figure 6.9.

The inversion gave distributions of phase velocity for four frequencies shown in Figure 6.10 a—d together with the original velocity distributions. One can see that, even though few data points of phase velocity were used, the resulting distributions of phase velocity reflect the principal features of the original distributions. One can thus expect that, also in the interpretation of real seismological data, a few observations of phase velocity are sufficient, when combined with a large sample of group velocity measurements, to derive fine detail in the distribution of phase velocity.

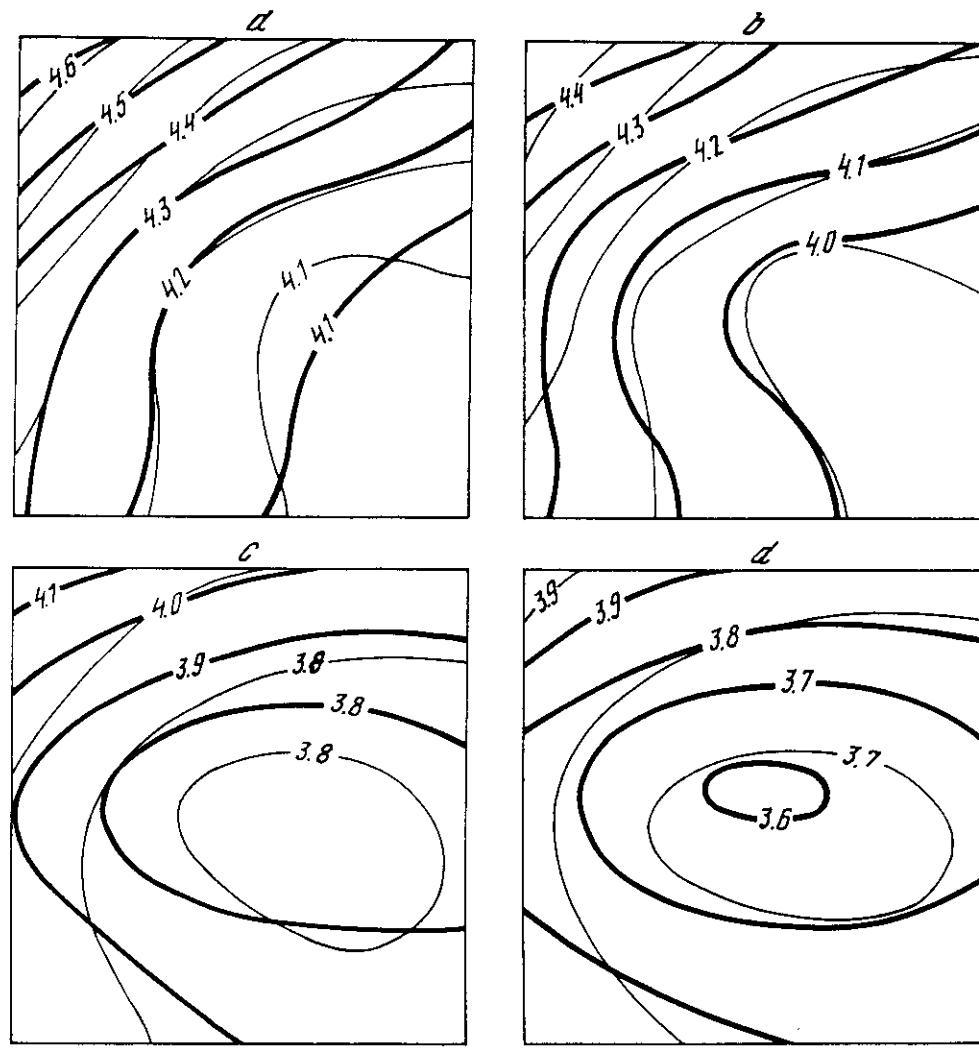

Fig. 6.10. Results of joint inversion of phase and group velocities. Solid lines show the initial phase velocity distributions, thin lines the results of inversion, for the frequencies: (a) $\omega = 0.21$ rad/sec, (b) $\omega = 0.28$ rad/sec, (c) $\omega = 0.42$ rad/sec, (d) $\omega = 0.50$ rad/sec.

5. Estimation of lateral velocity variations and anisotropy

The approach developed for determination of lateral velocity variations can be easily extended to estimation of azimuthal anisotropy.

For weak lateral velocity variations and weak anisotropy surface wave velocity may be approximated as follows:

$$V(x,y,\varphi) = V_0 + \delta V(x,y) + A(x,y)\cos 2\varphi + B(x,y)\sin 2\varphi.$$

Therefore unlike the isotropic media we have to determine three rather than one unknown functions: $\delta V(x, y), A(x, y)$ $B(x, y)$.

It is convenient to introduce the non-dimensional functions

$$m(x,y) = -\delta V / V_0,$$

$$a(x,y) = -A / V_0,$$

$$b(x,y) = -B / V_0.$$

Denote the angle between the i-th path and x-axis as φ_i. Then travel time along this path is

$$t_i = t_{0i} + \iint G_i(\mathbf{r})m(\mathbf{r})d\mathbf{r} + \cos 2\varphi_i \iint G_i(\mathbf{r})a(\mathbf{r})d\mathbf{r} + \sin 2\varphi_i \iint G_i(\mathbf{r})b(\mathbf{r})d\mathbf{r}.$$

Using the notation $G_{ci} = G_i \cos 2\varphi_i$, $G_{si} = G_i \sin 2\varphi_i$ we can write the mean square residual in the form:

$$S = (\mathbf{G}m + \mathbf{G}_c a + \mathbf{G}_s b - \delta t)^T \mathbf{R}_t^{-1} (\mathbf{G}m + \mathbf{G}_c a + \mathbf{G}_s b - \delta t).$$

The regularization functional can be taken analogously to the isotropic case, though in general the parameters of regularization should be different for lateral velocity variation and anisotropy:

$$\alpha \iint |\nabla m|^2 d\mathbf{r} + \beta \iint [|\nabla a|^2 + |\nabla b|^2]d\mathbf{r}.$$

As before the functions m, a and b are assumed to be finite at infinity. Then it can be shown that the following functional should be minimized:

$$(\mathbf{G}m + \mathbf{G}_c a + \mathbf{G}_s b - \delta\mathbf{t})^T \mathbf{R}_t^{-1} (\mathbf{G}m + \mathbf{G}_c a + \mathbf{G}_s b - \delta\mathbf{t}) + \alpha \iint m\nabla m d\mathbf{r} + \beta \iint [a\Delta a + b\Delta b]d\mathbf{r}. \quad (1)$$

Denote

$$\alpha^{-1}\mathbf{R}_t^{-1}(\mathbf{G}m + \mathbf{G}_c a + \mathbf{G}_s b - \delta\mathbf{t}) = \Lambda. \quad (2)$$

Then the minimization condition (1) is reduced to the following equations with respect to the unknown functions $m(\mathbf{r})$, $a(\mathbf{r})$ and $b(\mathbf{r})$:

$$\begin{aligned} \Delta m + \Lambda^T \mathbf{G} &= 0, \\ \Delta a + k\Lambda^T \mathbf{G}_c &= 0, \\ \Delta b + k\Lambda^T \mathbf{G}_s &= 0, \end{aligned} \quad (3)$$

where $k = \alpha/\beta$. . It follows from (3) :

$$m = -\sum \lambda_i \iint G_i \ln|\mathbf{r} - \mathbf{r}_i| d\mathbf{r} + C_1 = -\sum \lambda_i K_i(\mathbf{r}) + C_1,$$
$$a = -k\sum \lambda_i \cos 2\varphi_i \iint G_i \ln|\mathbf{r} - \mathbf{r}_i| d\mathbf{r} + C_2 = -\sum \lambda_{ci} K_i(\mathbf{r}) + C_2,$$
$$b = -k\sum \lambda_i \cos 2\varphi_i \iint G_i \ln|\mathbf{r} - \mathbf{r}_i| d\mathbf{r} + C_3 = -\sum \lambda_{si} K_i(\mathbf{r}) + C_3, \qquad (4)$$

where $\lambda_{ci} = k\lambda_i \cos 2\varphi_i$, $\lambda_{si} = k\lambda_i \sin 2\varphi_i$, and the functions $K_i(\mathbf{r})$ are determined by formulae (4).

Substitute these expressions into (2), and assume for simplicity that $\mathbf{R} = \mathbf{I}$ - unit matrix:

$$\lambda_j = -\frac{1}{\alpha}\Big[\sum \lambda_i Q_{ji} + C_1 t_{0j} + k\cos 2\varphi_j \sum \lambda_i \cos 2\varphi_i Q_{ji} + C_2 t_{0j} \cos 2\varphi_j +$$
$$+k\sin 2\varphi_j \sum \lambda_i \sin 2\varphi_i Q_{ji} + C_3 t_{0j} \sin 2\varphi_j + \delta t_j\Big],$$

where $Q_{ji} = \int\limits_{L_i} \int\limits_{L_j} \ln|\mathbf{r}_j - \mathbf{r}_i| ds_i ds_j / V_0^2$. Otherwise

$$(\tilde{\mathbf{Q}} + \alpha\mathbf{I})\boldsymbol{\Lambda} + C_1\mathbf{t}_0 + C_2\mathbf{t}_{0x} + C_3\mathbf{t}_{0y} = \delta\mathbf{t},$$

where

$$\tilde{Q}_{ji} = Q_{ji}[1 + k\cos 2(\varphi_i - \varphi_j)],$$
$$t_{0xi} = t_{0i}\cos 2\varphi_i,$$
$$t_{0yi} = t_{0i}\sin 2\varphi_i.$$

The conditions of finiteness of the functions m, a and b at infinity result in the three additional equations

$$\boldsymbol{\Lambda}^T\mathbf{t}_0 = 0,$$
$$\boldsymbol{\Lambda}^T\mathbf{t}_{0x} = 0,$$
$$\boldsymbol{\Lambda}^T\mathbf{t}_{0y} = 0.$$

Thus, the following system should be solved:

$$\begin{pmatrix} S_{11} & S_{12}\cdots\cdots S_{1N} & t_{01} & t_{0x1} & t_{0y1} \\ S_{21} & S_{22}\cdots\cdots S_{2N} & r_{02} & t_{0x2} & t_{0y2} \\ \vdots\vdots\vdots & & & & \\ S_{N1} & S_{N2}\cdots\cdots S_{NN} & t_{0N} & t_{0xN} & t_{0yN} \\ t_{01} & t_{02}\cdots\cdots t_{0N} & 0 & 0 & 0 \\ t_{0x1} & t_{0x2}\cdots\cdots t_{0xN} & 0 & 0 & 0 \\ t_{0y1} & t_{0y2}\cdots\cdots t_{0yN} & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_2 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} \delta t_1 \\ \delta t_2 \\ \vdots \\ \delta t_N \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad (5)$$

or $\mathbf{S}\Lambda^* = \delta t^*$, , where $\mathbf{S}$ is the matrix of the system (5); $\Lambda^{*T} = (\lambda_1, \lambda_2, ..., \lambda N, C_1, C_2, C_3$, δt^* is vector of the right-hand side of (5).

Finally (4) can be written as

$$m(r) = -\delta t^{*T} \mathbf{S}^{-1} \mathbf{K}(\mathbf{r}),$$

$$a(r) = -k\delta t^{*T} \mathbf{S}^{-1} \mathbf{K}_c(\mathbf{r}),$$

$$b(r) = -k\delta t^{*T} \mathbf{S}^{-1} \mathbf{K}_s(\mathbf{r}),$$

where

$$\mathbf{K}(\mathbf{r})^T = \{K_1(\mathbf{r}), ..., K_N(\mathbf{r}), 1, 0, 0\},$$
$$\mathbf{K}_c(\mathbf{r})^T = \{K_1(\mathbf{r})\cos 2\varphi_1, ..., K_N(\mathbf{r})\cos 2\varphi_N, 0, 1, 0\},$$
$$\mathbf{K}_s(\mathbf{r})^T = \{K_1(\mathbf{r})\sin 2\varphi_1, ..., K_N(\mathbf{r})\sin 2\varphi_N, 0, 0, 1\}.$$

Otherwise

$$m(\mathbf{r}) = -Q^T \delta t, \; a(\mathbf{r}) = -k\mathbf{P}^T \delta t, \; b(\mathbf{r}) = -k\mathbf{R}^T \delta t, \qquad (6)$$

where the vectors $\mathbf{Q}(\mathbf{r})$, $\mathbf{P}(\mathbf{r})$ and $\mathbf{R}(\mathbf{r})$ are determined by the first N elements of the vectors $\mathbf{S}^{-1}\mathbf{K}(\mathbf{r})$, $\mathbf{S}^{-1}\mathbf{K}_c(\mathbf{r})$, and $\mathbf{S}^{-1}\mathbf{K}_s(\mathbf{r})$, respectively.

Azimuth of the largest velocity is determined from a and b:

$$Az = \text{arctg}(a / b) / 2,$$

and the anisotropy coefficient defined as $(V_{max} - V_{min})/V_{av}$ is

$$\kappa = 2\sqrt{a^2 + b^2}.$$

Estimation of the resolution

The resolution can be estimated only for those quantities, which are involved linearly to the equation for the velocity, i.e for
We may estimate the resolution of these functions by radii of the averaging areas as in the isotropic case.

For accurate data

$$\delta t = \mathbf{G}m + \mathbf{G}_c a + \mathbf{G}_s b. \qquad (7)$$

Substituting (7) into (6) we obtain

$$m(\mathbf{r}) = -\mathbf{Q}^T(\mathbf{G}m + \mathbf{G}_c a + \mathbf{G}_s b) = M_m m + M_a a + M_b b,$$

$$a(\mathbf{r}) = -\mathbf{P}^T(\mathbf{G}m + \mathbf{G}_c a + \mathbf{G}_s b) = A_m m + A_a a + A_b b,$$

$$b(\mathbf{r}) = -\mathbf{R}^T(\mathbf{G}m + \mathbf{G}_c a + \mathbf{G}_s b) = B_m m + B_a a + B_b b.$$

The notation of the type $M_m\mathbf{m}$ is of the following meaning:

$$M_m\mathbf{m} = \iint M_m(\mathbf{r},\mathbf{r}')m(\mathbf{r}')d\mathbf{r}'.$$

In order that the solution would reflect the real distributions in the best way, the functions $M_m(\mathbf{r}, \mathbf{r}')$, $A_a(\mathbf{r}, \mathbf{r}')$, $B_b(\mathbf{r}, \mathbf{r}')$ should be close to the delta-functions, and the other kernels close to zero.

Analogously to the isotropic case we can define the functionals

$$\begin{aligned}
s_m^2(\mathbf{r}) &= \int|\mathbf{E}_m^m(\mathbf{r},\mathbf{r}') - \varepsilon(\mathbf{r},\mathbf{r}')|^2\,d\mathbf{r}' + k\int|\mathbf{E}_m^a(\mathbf{r},\mathbf{r}')|^2\,d\mathbf{r}' + k\int|\mathbf{E}_m^b(\mathbf{r},\mathbf{r}')|^2\,d\mathbf{r}',\\
s_a^2(\mathbf{r}) &= \int|\mathbf{E}_a^m(\mathbf{r},\mathbf{r}')|^2\,d\mathbf{r}' + k\int|\mathbf{E}_a^a(\mathbf{r},\mathbf{r}') - \varepsilon(\mathbf{r},\mathbf{r}')|^2\,d\mathbf{r}' + k\int|\mathbf{E}_a^b(\mathbf{r},\mathbf{r}')|^2\,d\mathbf{r}',\\
s_b^2(\mathbf{r}) &= \int|\mathbf{E}_b^m(\mathbf{r},\mathbf{r}')|^2\,d\mathbf{r}' + k\int|\mathbf{E}_b^a(\mathbf{r},\mathbf{r}')|^2\,d\mathbf{r}' + k\int|\mathbf{E}_b^b(\mathbf{r},\mathbf{r}') - \varepsilon(\mathbf{r},\mathbf{r}')|^2\,d\mathbf{r}',
\end{aligned} \tag{8}$$

where $\operatorname{div} E_m^q = M_m^q$, $\operatorname{div} E_a^q = A_a^q$, $\operatorname{div} E_b^q = B_b^q$; the kernels M_m, A_a, B_b are normalized:

$$\int M_m(\mathbf{r},\mathbf{r}')d\mathbf{r}' = 1,$$

$$\int A_a(\mathbf{r},\mathbf{r}')d\mathbf{r}' = 1,$$

$$\int B_b(\mathbf{r},\mathbf{r}')d\mathbf{r}' = 1,$$

and the integrals of other kernels are equal to zero.

From these conditions we obtain

$$\mathbf{Q}^T\mathbf{t}_0 = 1, \quad \mathbf{P}^T\mathbf{t}_{0x} = 1, \quad \mathbf{R}^T\mathbf{t}_{0y} = 1,$$

$$\mathbf{Q}^T\mathbf{t}_{0x} = 0, \quad \mathbf{P}^T\mathbf{t}_0 = 0, \quad \mathbf{R}^T\mathbf{t}_0 = 0,$$

$$\mathbf{Q}^T\mathbf{t}_{0y} = 0, \quad \mathbf{P}^T\mathbf{t}_{0y} = 0, \quad \mathbf{R}^T\mathbf{t}_{0x} = 0.$$

After transformation of the functionals (8) and equating their values to those corresponding to the kernels, which differ from zero within the circles with radii R_m, R_a, R_b respectively, we obtain the following formulae for the effective radii of the averaging areas

$$R_m = \exp(3/4 - \mathbf{Q}^T\tilde{\mathbf{S}}\mathbf{Q} + 2\mathbf{K}_m^T\mathbf{Q}),$$

$$R_a = \exp[3/4 + (-\mathbf{P}^T\tilde{\mathbf{S}}\mathbf{P} + 2\mathbf{K}_a^T\mathbf{P})/k],$$

$$R_b = \exp[3/4 + (-\mathbf{R}^T\tilde{\mathbf{S}}\mathbf{R} + 2\mathbf{K}_b^T\mathbf{R})/k].$$

References

Ditmar P.G., Yanovskaya T.B. 1987. *Extension of the Backus-Gilbert method for estimation of lateral variations of surface wave velocity* . Izvestia AN SSSR, Fizika Zemli, 6, 30-40 (in Russian)

Nolet G., 1990. *Partitioned waveform inversion and two-dimensional structure under the network of autnomously recording seismographs.* J.Geophys.Res., 95, No.6, 8499-8512.

Snieder R., 1988. *Large-scale waveform inversions of surface waves for lateral heterogeneity. 2. Application to surface waves in Europe and Mediterranean.* J.Geophys.Res., 93, No. B10, 12067-12080.

Tarantola A., Nersessian A. *Three-dimensional inversion without blocks,* Geophys.J.Roy.astr.Soc., 1984, V.76, P. 299-306.

Yanovskaya T.B.,Maaz R., Ditmar P.G., Neunhofer H. 1988, *A method for joint interpretation of the phase and group surface wave velocities to estimate lateral variations of the Earth's structure.* Phys.Earth and Planet.Inter., 51, 59-67.

Yanovskaya T.B., Ditmar P.G. *Smoothness criteria in surface wave tomography.* Geophys.J.Int., 1990, V.102, P. 63-72

PAGEOPH, Vol. 147, No. 3 (1996)
0033-4553/96/000000-00$1.50 + 0.20/0

Ray Tomography Based on Azimuthal Anomalies

T. B. YANOVSKAYA[1]

Abstract. A method of estimating the lateral velocity variations in the 2D case using the data on deviations of wave paths from straight lines (or great circle paths in the spherical case) is proposed. The method is designed for interpretation of azimuthal anomalies of surface waves which contain information on lateral variations of phase velocities supplementary to that obtained from travel-time data in traditional surface wave tomography.

In the particular 2D case, when the starting velocity is constant (c_0) and velocity perturbations $\delta c(x, y)$ are sufficiently smooth, a relationship between azimuthal anomaly $\delta\alpha$ and velocity perturbations $\delta c(x, y)$ can be obtained by approximate integration of the ray tracing system, which leads to the following functional:

$$\delta\alpha = \int_0^L \frac{s(\nabla m, n_0)}{L}\, ds,$$

where $m(x, y) = \delta c(x, y)/c_0$, L is the length of the ray, n_0 is a unit vector perpendicular to the ray in the starting model, integration being performed from the source to the receiver. This formula is valid for both plane and spherical cases. Numerical testing proves that for a velocity perturbation which does not exceed 10%, this approximation is fairly good. Lateral variations of surface wave velocities satisfy these assumptions. Therefore this functional may be used in surface wave tomography.

For the determination of $m(x, y)$ from a set of $\delta\alpha_k$ corresponding to different wave paths, the solution is represented as a series in basis functions, which are constructed using the criterion of smoothness of the solution proposed by TARANTOLA and NERSESSIAN (1984) for time-delay tomography problems. Numerical testing demonstrates the efficiency of the tomography method.

The method is applied to the reconstruction of lateral variations of Rayleigh wave phase velocities in the Carpathian-Balkan region. The variations of phase velocities obtained from data on azimuthal anomalies are found to be correlated with group-velocity variations obtained from travel-time data.

Key words: Surface waves, phase velocities, azimuthal anomalies, seismic tomography.

1. *Introduction*

The data which are widely used in tomographic reconstruction of lateral velocity variations are travel times of seismic waves. This problem can be easily linearized: the time delay with respect to a properly chosen starting velocity model is represented as a linear functional of the unknown velocity variations. Thus the tomography problem is reduced to a system of linear equations. Alternative data,

[1]Institute of Physics, Sankt-Petersburg State University, Petrodvoretz, Sankt-Petersburg 198904, Russia.

which also depend on velocity variations and can be easily obtained from seismological observations, are polarization anomalies, which are related to anomalies of direction of wave propagation.

Data on surface wave polarization allow the azimuth of the wave arriving at a station to be determined. Accordingly the *azimuthal anomaly*, which is a deviation of the observed azimuth from that corresponding to the great circle path, contains information about the lateral variation of surface-wave phase velocities. Thus it seems to be expedient to incorporate azimuth anomalies in surface wave tomography.

Observations of surface waves indicate the existence of prominent azimuth anomalies in some cases (LANDER, 1984; LERNER-LAM and PARK, 1989; NESTEROV and YANOVSKAYA, 1988; LEVSHIN *et al.*, 1994; LASKE *et al.*, 1994). However, these anomalies have so far been interpreted practically only qualitatively. The reason for which azimuthal anomalies are not used in tomographic reconstruction (separately or jointly with travel-time data) is the difficulty in determining the linear functional, which relates velocity variations and azimuthal anomalies similar to that for time delays. Recently HU and MENKE (1992) proposed a formalism for calculating the matrix for transforming model parameters to polarization anomalies and applied it to the determination of the P-wave velocity structure in southern California from polarization data (Hu *et al.*, 1994). In this approach polarization anomalies are related to model parameters rather than to velocity variations directly, so that this approach requires the preliminary parametrization of the velocity model. However, when a set of data is rather poor—and this situation is customary surface wave tomography—in tomographic studies it is more expedient to use a method which is not based on *a priori* parametrization (TARANTOLA and NERSESSIAN, 1984; YANOVSKAYA and DITMAR, 1990). However in this case the data should be expressed in the form of a linear functional of the unknown velocity variation.

The present paper demonstrates how to simplify the relationship between azimuth anomalies and lateral phase velocity variations in the 2D case of the velocity in the starting model may be assumed to be constant. This case is applicable to surface wave data. A method of inverting the azimuth anomalies to lateral velocity variations is also proposed.

2. *Approximate Relationship between Velocity Variations and Azimuthal Anomalies for Constant Initial Velocity*

We consider a 2D model, the velocity in the starting model being constant. The surface wave velocity corresponding to a fixed period satisfies this assumption: in fact, lateral variations of phase velocities are small, and in the first approximation the surface waves propagate along great circle paths, or along straight lines in the

plane case. In general the problem is formulated as follows:

to derive a relationship between azimuthal anomaly δx and lateral variation of phase velocity $\delta c(\mathbf{x}) = c(\mathbf{x}) - c_0$ in linear approximation,

where x signifies the coordinates on the surface (plane or spherical).

Plane Case

It is convenient to express the ray tracing system in the form

$$\frac{d\mathbf{x}}{dq} = S\mathbf{t} \tag{1a}$$

$$\frac{d\mathbf{t}}{dq} = -S\mathbf{n}\frac{(\nabla c, \mathbf{n})}{c} \tag{1b}$$

where $\mathbf{t}$ and $\mathbf{n}$ are unit vectors tangent and orthogonal to the ray, respectively, S is the length of the ray, q a parameter varying from 0 to 1, so that $ds = S\,dq$, where ds is an element of the ray length. This parameterization was introduced by JULIAN and GUBBINS (1977) for simple mapping from initial ray to perturbed ray, and it is useful in studies dealing with ray perturbation (e.g., SNIEDER and SPENCER, 1993).

Assuming $\nabla c_0 = 0$, the system for the variations $\delta\mathbf{x}$ and $\delta\mathbf{t}$ can be expressed as follows

$$\frac{d\delta\mathbf{x}}{dq} = \mathbf{t}_0\,\delta S + S_0\,\delta\mathbf{t} \tag{2a}$$

$$\frac{d\delta\mathbf{t}}{dq} = -S_0\frac{(\nabla\delta c, \mathbf{n}_0)}{c_0}\mathbf{n}_0. \tag{2b}$$

The value of $\nabla\delta c$ should be taken along the ray corresponding to the model $c(x, y)$, i.e., at points $\mathbf{x} = \mathbf{x}_0 + \delta\mathbf{x}$, so that

$$\nabla\delta c(\mathbf{x}) = \nabla\delta c(\mathbf{x}_0) + \left(\frac{\partial\nabla\,\delta c}{\partial\mathbf{x}}\right)_{\mathbf{x}_0}\delta\mathbf{x}. \tag{3}$$

In the case of smooth velocity variations, the last term on the r.h.s. of (3) is of the second order of smallness. Indeed, if the correlation length of the heterogeneities is L, this term is of the order of $\delta c|\delta\mathbf{x}|/L^2$. Thus this term may be neglected if L is not too small, i.e., in the case of smooth velocity variations.

Substituting $\nabla\delta c(\mathbf{x})$ in (2b) by $\nabla\delta c(\mathbf{x}_0)$ and integrating this equation from the receiver to the source we obtain:

$$\delta\mathbf{t}(q) = \delta\mathbf{t}_r - S_0\int_0^q \frac{(\nabla\delta c, \mathbf{n}_0)}{c_0}\mathbf{n}_0\,dq' \tag{4}$$

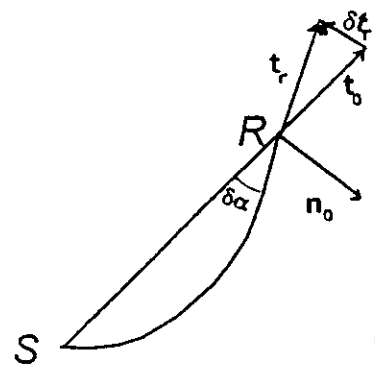

Figure 1
Scheme of rays, unit vectors $\mathbf{t}_0$, $\mathbf{n}_0$ and azimuthal anomaly $\delta\alpha$.

where $\delta\mathbf{t}_r$ is the variation of unit vector t at the receiver. Vector $\delta\mathbf{t}_r$ as well as unit vectors $\mathbf{t}_0$ and $\mathbf{n}_0$ are shown schematically in Figure 1.

Substituting (4) into equation (2a) and integrating again from the receiver to the source, i.e., from $q = 0$ to $q = 1$, yields

$$\delta\mathbf{x} = \mathbf{t}_0\,\delta S + S_0 \int_0^1 \delta\mathbf{t}_r\,dq - S_0^2 \int_0^1 dq \int_0^q \frac{(\nabla\delta c, \mathbf{n}_0)}{c_0}\,\mathbf{n}_0\,dq'. \tag{5}$$

Taking into account that $\delta\mathbf{x} = 0$ at the source, and integrating along the undisturbed ray ($ds = S_0\,dq$) we obtain

$$S_0\,\delta\mathbf{t}_r = \int_0^{S_0} ds \int_0^s \frac{(\nabla\delta c, \mathbf{n}_0)}{c_0}\,\mathbf{n}_0\,ds' - \mathbf{t}_0\,\delta S. \tag{6}$$

The second term on the r.h.s. of (6) should be at least of the second order of smallness: the variation of the unit vector should be orthogonal to the vector itself, and in linear approximation

$$(\delta\mathbf{t}_r, \mathbf{t}_0) \approx 0, \quad \delta\alpha \cong (\delta\mathbf{t}_r, \mathbf{n}_0),$$

where $\delta\alpha$ is the deviation of the azimuth at the receiver (see Fig. 1). Consequently,

$$S_0\,\delta\alpha = \int_0^{S_0} ds \int_0^s \frac{(\nabla\delta c, \mathbf{n}_0)}{c_0}\,ds'. \tag{7}$$

The double integral in (7) can be transformed to a single integral by changing the order of integration:

$$\int_0^{S_0} ds \int_0^s f(s')\,ds' = \int_0^{S_0} f(s')\,ds' \int_s^{S_0} ds = \int_0^{S_0} (S_0 - s')f(s')\,ds'.$$

Thus

$$\delta\alpha = \int_0^S (S_0 - s)\frac{(\nabla\delta c, \mathbf{n}_0)}{S_0 c_0}\,ds. \tag{8}$$

Assuming the opposite direction of integration, from the source to the receiver, we obtain

$$\delta\alpha = \int_0^{S_0} \frac{s}{S_0 c_0}(\nabla\delta c, \mathbf{n}_0)\,ds = \int_0^{S_0} \frac{s}{S_0 c_0}\left(\frac{\partial\delta c}{\partial x} n_{0x} + \frac{\partial\delta c}{\partial y} n_{0y}\right) ds. \tag{9}$$

Thus the azimuthal anomaly is expressed in the form of a linear functional of the spatial derivatives of $\delta c(x, y)$.

Equation (9) shows that the azimuthal anomaly is affected by the component of the gradient of the velocity variation orthogonal to the ray. It follows from the facts that a constant perturbation of the velocity has no effect on the shape of the ray; and that the shape of the ray remains unchanged if the velocity perturbation varies only along the ray. This agrees with the conclusion drawn by HU *et al.* (1994) and FARRA and LE BÉGAT (1995) regarding the sensitivity of polarization data: they are most sensitive to local heterogeneities, whereas travel-time anomalies are affected by the long waelength components of velocity variations: in the former case the velocity gradient is larger. Also equation (9) clearly supports the conclusion (HU *et al.*, 1994) that the velocity perturbation in the vicinity of the receiver has a stronger effect on the polarization anomaly than that in the vicinity of the source: the contribution of $(\nabla\delta c, \mathbf{n}_0)$ to the azimuthal anomaly is proportional to the distance s from the source.

In the tomography problem it is convenient to represent the functional as an integral over the 2D area. As is usually adopted in time delay tomography, we introduce a dimensionless unknown function $m(x) = \delta c(x)/c_0$. After transforming the Cartesian coordinate system such that the origin is located at the source and the x axis directed along the ray, (9) can be modified to read:

$$\delta\alpha = \iint_\Sigma (\nabla m, \mathbf{F})\,d\mathbf{x} \tag{10}$$

where $\mathbf{F}(x, y) = x\delta(y)\{H(x) - H(x - S_0)\}\mathbf{e}y$, $\delta(y)$ is the Dirac function, and $H(x)$ is the Heaviside function.

Spherical Case

As opposed to travel times, azimuths are not invariant to transformation of a plane to a spherical surface. Therefore, it is necessary to derive a relationship between the azimuthal anomaly and perturbation of the velocity gradient directly on the sphere.

The ray tracing equations on a spherical surface of radius R are:

$$d\mathbf{x}/ds = \mathbf{t}, \quad d\mathbf{t}/ds = -\frac{(\nabla c, \mathbf{n})}{c}\mathbf{n} - \frac{\mathbf{e}_R}{R}$$

where $\mathbf{e}_R$ is the unit vector along the radius. (Now we may differentiate with respect to s rather than to q: as was shown above, the variation of ray length δS has no effect on the azimuthal anomaly.

The equations for variations of x and t read as follows:

$$\frac{d\delta\mathbf{x}}{ds} = \delta\mathbf{t}, \quad \frac{d\delta\mathbf{t}}{ds} = -\frac{(\nabla\delta c, \mathbf{n}_0)}{c_0} - \frac{\delta\mathbf{e}_R}{R}. \tag{11}$$

The variation of unit vector $\mathbf{e}_R$ can be expressed approximately in terms of the deviation of the ray from the great circle $\delta\mathbf{x}$: $\delta\mathbf{e}_R = \delta\mathbf{x}/R$. In the plane case system (11) thus differs from system (2a,b) by the additional term $-\delta\mathbf{x}/R^2$ in the second equation. This term may be neglected if it is of the same order as other terms being neglected.

In the plane case we assume that $\nabla\delta c(\mathbf{x}) \cong \nabla\delta c(\mathbf{x}_0)$, neglecting the term

$$\left[c_0^{-1}\left(\frac{\partial\nabla\,\delta c}{\partial\mathbf{x}}\right)_{\mathbf{x}=\mathbf{x}_0}\delta\mathbf{x}, \mathbf{n}_0\right]\mathbf{n}_0$$

in equation (2b). If the correlation length of heterogeneities is L, it is easy to see that this term is of the order of $(\delta c|\delta\mathbf{x}|)/c_0L^2$. The additional term in the second equation (11) would be of the same order if $\delta c/c_0 \sim L^2/R^2$. This means that if the dimension of heterogeneities is not too large, this term may also be neglected. Thus the approximate relationship between the azimuth anomaly and gradient of velocity variation should be the same as in the plane case (formula (9)).

3. *Numerical Testing*

The validity of eq. (9) was tested numerically. Obviously the larger the second derivatives of the velocity variations, the larger the errors due to the neglect of the second term on the r.h.s. of (3). For the velocity model shown in Figure 2a, the rays shown schematically (as straight lines) in Figure 2b have been calculated by integrating the ray tracing system. The calculated rays allow 'exact' azimuth anomalies to be determined. The azimuthal anomalies were then calculated using eq. (9). The exact and approximate azimuthal anomalies are shown in Figure 3. The errors due to the adopted approximation are not too large and, therefore, the approximate formula for calculating the azimuthal anomalies is acceptable for models with smooth velocity variations as in the case considered.

Numerical testing performed for many different velocity models confirms this conclusion.

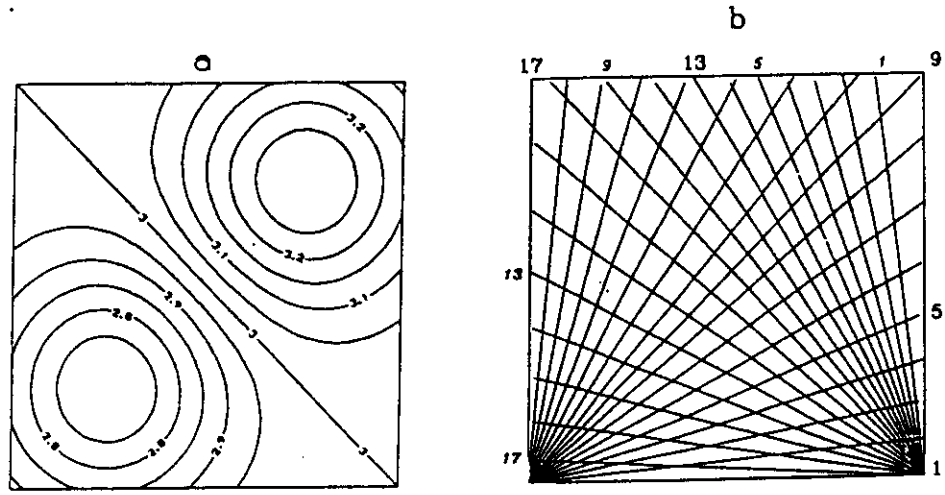

Figure 2
Velocity model (2a) and pattern of rays (2b) used for testing the validity of the approximate formula for azimuth anomalies. The rays are numbered from 1 to 17 for each set of rays.

4. *A Method for Azimuthal Tomography on a Plane*

Once the azimuthal anomaly is represented in the form of a linear functional, the traditional approaches developed in time-delay tomography may be used for reconstructing the velocity variations from azimuthal anomalies. However, this

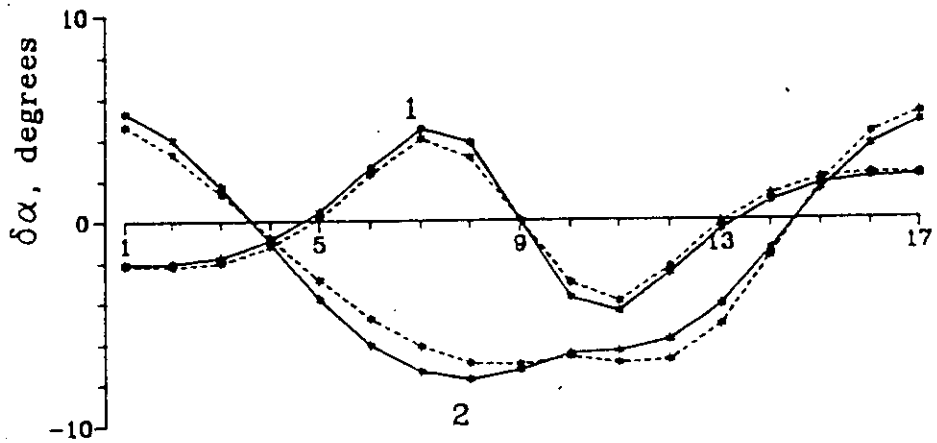

Figure 3
Exact (solid lines) and approximate (dashed lines) azimuth anomalies for the rays shown in Figure 2b. The numbers of the rays are shown along the x axis. Curves 1 and 2 correspond to the sets of rays from sources in the left and in the right corners, respectively.

problem (later referred to as 'azimuthal tomography') has specific features connected with the data being related to the velocity gradient rather than to the velocity itself. Therefore, some standard methods, such as discretization of the model, are invalid in solving this problem. It is clear that the unknown model should be sufficiently smooth in order to eliminate large values and jumps of the gradient.

A general approach for solving the tomography problem is to represent the unknown function as a set in some basis functions $\psi_i(\mathbf{x})$:

$$m(\mathbf{x}) = \sum_i \lambda_i \psi_i(\mathbf{x}) \tag{12}$$

and thereby to reduce the problem to a linear system in the unknown coefficients λ_i. Obviously, a key point in such an approach is the suitable choice of the basis functions. The best way is to construct the basis functions using some reasonable *a priori* assumptions concerning the unknown function, as proposed by TARANTOLA and NERSESSIAN (1984), or YANOVSKAYA and DITMAR (1990), because in this case the basis functions would be in agreement with the data set. As mentioned above, in azimuthal tomography it is necessary to assume the function to be sufficiently smooth.

The smoothness criterion proposed by TARANTOLA and NERSESSIAN (1984) ensures the smoothness of function $m(x)$ as well as of all its derivatives, as was shown by YANOVSKAYA and DITMAR (1990). Therefore this criterion is suitable for the problem of azimuthal tomography, and this is the reason why it was chosen in the present study.

For exact data this criterion is reduced to the minimization of the functional

$$\iint m(\mathbf{x}) C^{-1}(\mathbf{x}, \mathbf{x}') m(\mathbf{x}')\, d\mathbf{x}\, d\mathbf{x}' \tag{13}$$

where $C(\mathbf{x}, \mathbf{x}')$ is an *a priori* covariance function of the model.

To determine a solution satisfying the minimization of (13) under constraints (10) it is convenient to represent (10) as a functional of $m(\mathbf{x})$ rather than of its gradient. This can be done by applying Green's formula:

$$\int_\Sigma (\nabla m, \mathbf{F})\, d\mathbf{x} = \int_{C_\Sigma} m F_n\, dl - \int_\Sigma m \operatorname{div} \mathbf{F}\, d\mathbf{x}$$

where C_Σ is the contour of area Σ, and F_n is the component of $\mathbf{F}$ normal to the contour. If all the rays are inside Σ, then, according to the definition of $\mathbf{F}$, $F_n = 0$ along the contour, and consequently,

$$\int_\Sigma (\nabla m, \mathbf{F})\, d\mathbf{x} = -\int_\Sigma m(\mathbf{x}) \operatorname{div} \mathbf{F}(\mathbf{x})\, d\mathbf{x}. \tag{14}$$

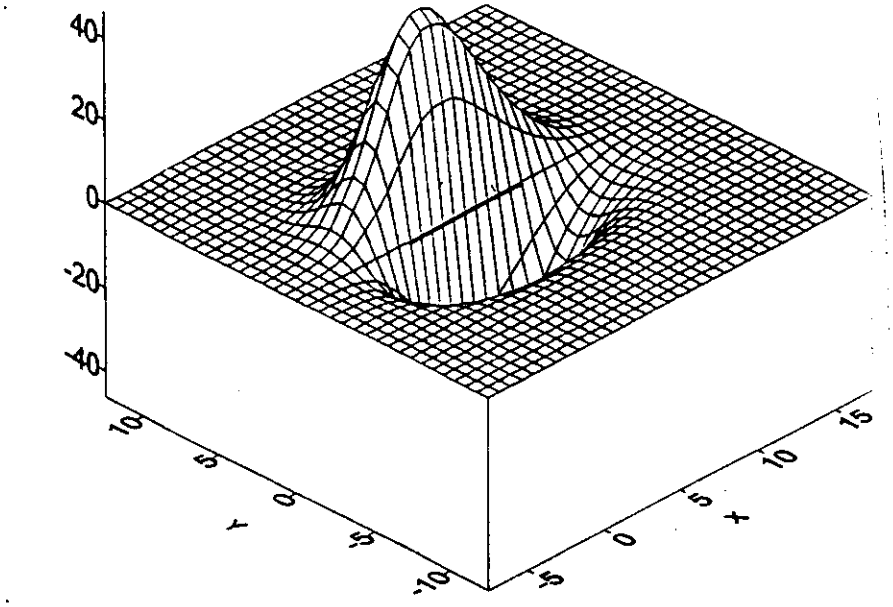

Figure 4
Example of the basis function (18) for the ray represented by the bold line.

Now the problem has been formulated as the minimization of (13) under constraints

$$\int_{\Sigma} m(\mathbf{x})\ \mathrm{div}\ \mathbf{F}_i(\mathbf{x})\, d\mathbf{x} = -\delta\alpha_i, \tag{15}$$

where $i = 1, 2, \ldots, N$, and N is the number of observations.

Analogously to TARANTOLA and NERSESSIAN (1984) it may be shown that the solution of this problem is:

$$m(\mathbf{x}) = \sum_{i=1}^{N} \lambda_i \int_{\Sigma} C(\mathbf{x}, \mathbf{x}')\ \mathrm{div}\ \mathbf{F}_i(\mathbf{x}')\, d\mathbf{x}' \tag{16}$$

and consequently, the basis functions are as follows:

$$\psi_i(\mathbf{x}) = \int_{\Sigma} C(\mathbf{x}, \mathbf{x}')\ \mathrm{div}\ \mathbf{F}_i(x')\, d\mathbf{x}'. \tag{17}$$

If the *a priori* covariance function of the model is assumed in the form

$$C(\mathbf{x}, \mathbf{x}') = \sigma^2 \exp\left(-\frac{|\mathbf{x} - \mathbf{x}'|^2}{2L^2}\right)$$

where L is the correlation length, the basis functions (17) may then be expressed in coordinates x, y connected with the ray, as indicated above, in explicit form up to

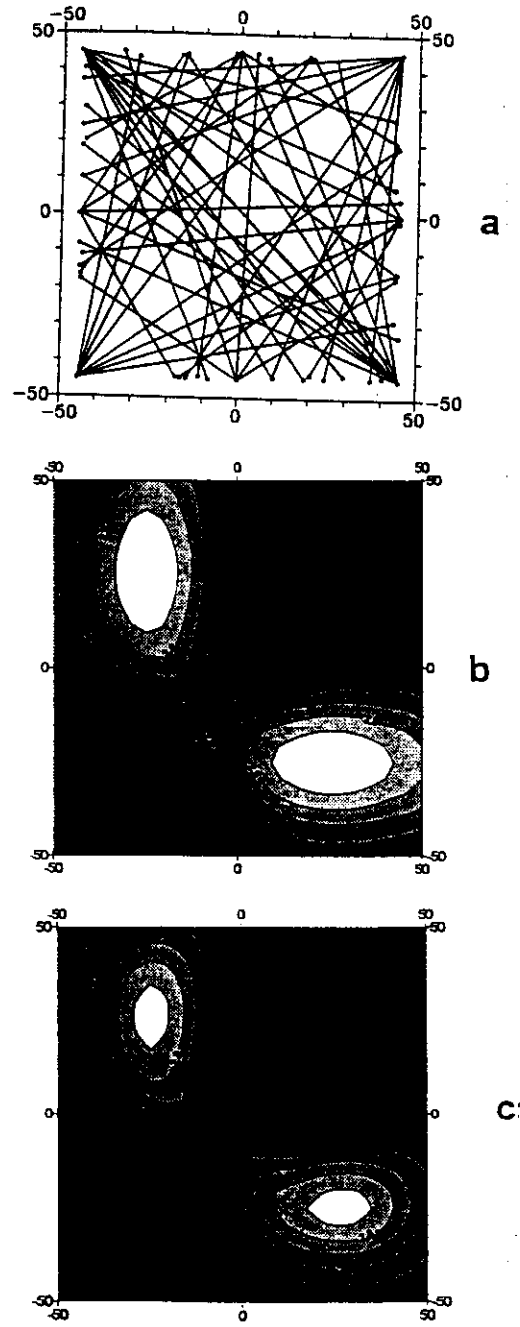

Figure 5
Solution of the tomography problem—model example, plane case. a—pattern of rays, b—velocity model, c—solution.

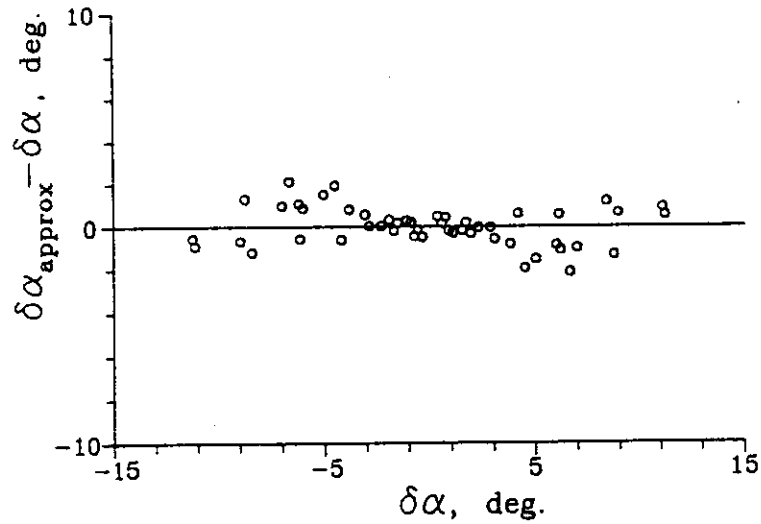

Figure 6
Deviation of approximate from exact azimuths as a function of azimuth anomalies for the rays shown in Figure 5a.

a constant factor σ^2:

$$\psi_i(x, y) = \left\{\sqrt{\frac{\pi}{2}}(S_{0i} - x)\left[\operatorname{erf}\left(\frac{S_{0i} - x}{L\sqrt{2}}\right) + \operatorname{erf}\left(\frac{x}{L\sqrt{2}}\right)\right] + L\left[\exp\left(-\frac{(S_{0i} - x)^2}{2L^2}\right) - \exp\left(-\frac{x^2}{2L^2}\right)\right]\right\}\frac{y}{LS_{0i}}\exp\left(-\frac{y^2}{2L^2}\right). \tag{18}$$

The behavior of the basis function is shown in Figure 4.

Inserting (12) into (10) we obtain a linear system of equations for the vector Λ of the unknown coefficients λ_i;

$$\mathrm{A}\Lambda = \delta\alpha$$

where $\delta\alpha$ is the vector of data $\delta\alpha_i$, and

$$A_{ij} = \int_0^{S_0} \frac{\partial\psi_j(x, y)}{\partial n_{0i}} \frac{s}{S_{0i}}\, ds. \tag{19}$$

If the data are inaccurate, and the statistical properties of the data errors are described by covariance matrix **R**, the solution should be determined by minimizing functional

$$\iint m(\mathrm{x})C^{-1}(\mathrm{x}, \mathrm{x}')m(\mathrm{x}')\, d\mathrm{x}\, d\mathrm{x}' - \sum_{ij}\left(\delta\alpha_i - \int m \operatorname{div} \mathbf{F}_i\, d\mathrm{x}\right)R_{ij}^{-1}\left(\delta\alpha_j + \int m \operatorname{div} \mathbf{F}_j\, d\mathrm{x}\right), \tag{20}$$

where R_{ij}^{-1} are elements of matrix $\mathbf{R}^{-1}$. It is easy to see that coefficients λ_i are determined from the system

$$(\mathbf{A} + \mathbf{R}/\sigma^2)\Lambda = \delta\alpha,$$

where $\mathbf{A}$ is defined by (19). If the data are independent and have one and the same standard error σ_d, this equation becomes:

$$(\mathbf{A} + \gamma\mathbf{I})\Lambda = \delta\alpha, \tag{21}$$

where $\mathbf{I}$ is the unit matrix, and $\gamma = \sigma_d^2/\sigma^2$ may be regarded as the regularization parameter.

Figure 5 demonstrates the reconstruction of the model velocity distribution using the above technique. The azimuthal anomalies have been calculated by integrating the ray tracing system along the rays drawn schematically in Figure 5a for the velocity model shown in Figure 5b.

Calculations for different values of correlation length L have shown that the optimal value is that corresponding to the mean distance between paths, which is estimated as $\sqrt{S/N}$, where S is the area covered by the paths, and N is the number of paths. For larger L the solution is too smooth, and for smaller L it becomes concentrated in the narrow vicinity of the paths. In this example L was taken to be equal to 12.

The agreement of the result of the inversion (Fig. 5c) with the initial model seems to be striking, if we take into account that the tomographic solution is based on the approximate relationship between the azimuthal anomaly and velocity variation.

The errors in azimuths due to approximation are shown in Figure 6 for the rays in Figure 5a. Though they are large enough (about 5–10%), they do not display any systematic behavior. Perhaps this is the reason why the tomographic reconstruction is so similar to the initial model: the errors due to the use of an approximate formula may be regarded as random errors of observations, which are suppressed when a large body of data is used.

5. *Azimuthal Tomography on a Spherical Surface*

Since the relationship between the azimuthal anomaly and velocity gradient in a spherical case is identical to a plane case, the tomography technique is analogous to that described in the previous section. A solution is represented as a set in basis functions (12), which corresponds to the smoothness criterion (13). However, in the spherical case it is impossible to represent the basis function in terms of elementary functions: it can only be expressed in the form of a double integral, which is inconvenient for numerical computation, the more so, as an additional numerical integration is required to calculate matrix A. Although in principle, a solution may

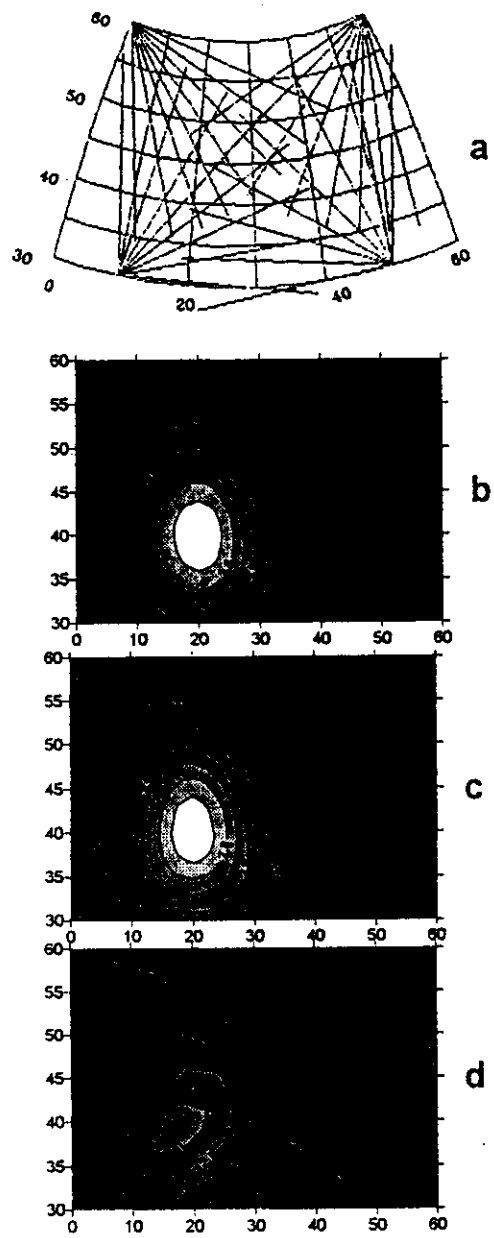

Figure 7

Solution of the tomography problem—model example, spherical case. a—pattern of rays, b—velocity model, c—solution for the case when 4 sources are placed in the corners of the area, d—solution for the case when sources and receivers are exchanged.

be constructed as a set in *any* basis functions. Therefore, for simplicity we may take basis functions (18), assuming x and y to be the distance along the ray and distance along the great circle orthogonal in the ray. Such basis functions differ insignificantly from those obtained from the criterion (13). Of course a solution expanded in such basis functions does not strictly minimize (13) or (20), but it should be adequate to the data set, which is not the case if some arbitrary basis functions, e.g., spherical functions, are assumed.

Figure 7 shows a numerical example of the velocity reconstruction in the spherical case when basis functions (18) were used. This figure also demonstrates the sensitivity of azimuthal anomalies to the location of velocity heterogeneities. As was mentioned in Section 2, the azimuthal anomalies are more sensitive to local heterogeneities in the vicinity of receivers rather than to those near sources. Figure 7c shows the tomographic reconstruction of the velocity when the sources are located in the corners of the area, where the velocity varies smoothly and the receivers are located near strong velocity perturbations. In this case the solution reflects the peculiarities of the model (Fig. 7b) well enough. However, if the receivers and sources are exchanged, the sensitivity of the data to the local heterogeneities drops, and the result of the tomographic inversion becomes much less satisfactory (Fig. 7d).

6. *Laterial Variations of Rayleigh Wave Phase Velocities in the Carpathian-Balkan Region from Azimuth Anomalies*

Unfortunately, a sufficient amount of reliable data on azimuthal anomalies of surface waves applicable to tomographic inversion is not available. Therefore, we must restrict the study to an example of applying the above technique to data which are not sufficiently accurate, over paths which do not cover the area under investigation densely. Such data on azimuthal anomalies of Rayleigh waves for periods of 10–25 s in the Carpathian-Balkan region have been obtained by NESTEROV and YANOVSKAYA (1988) using spectral polarization analysis. The wave paths for which the polarization data are available are shown in Figure 8. The rays are drawn schematically as circular arcs with the observed azimuthal anomalies at the stations.

The technique described above is employed to reconstruct the lateral phase velocity variations for periods $T = 10$ s and $T = 25$ s.

The solution is determined in the central part of the region, covered most densely by wave paths (it is contoured by the dashed line in Fig. 8). Since the data on azimuthal anomalies yield only velocity variations, the mean velocities have been assumed *a priori*: 3 km/s and 3.3 km/s for $T = 10$ s and 25 s, respectively. Correlation length L is taken to be equal to 300 km according to considerations described in Section 4, though the result differs insignificantly from those of $L = 200$ km and $L = 500$ km.

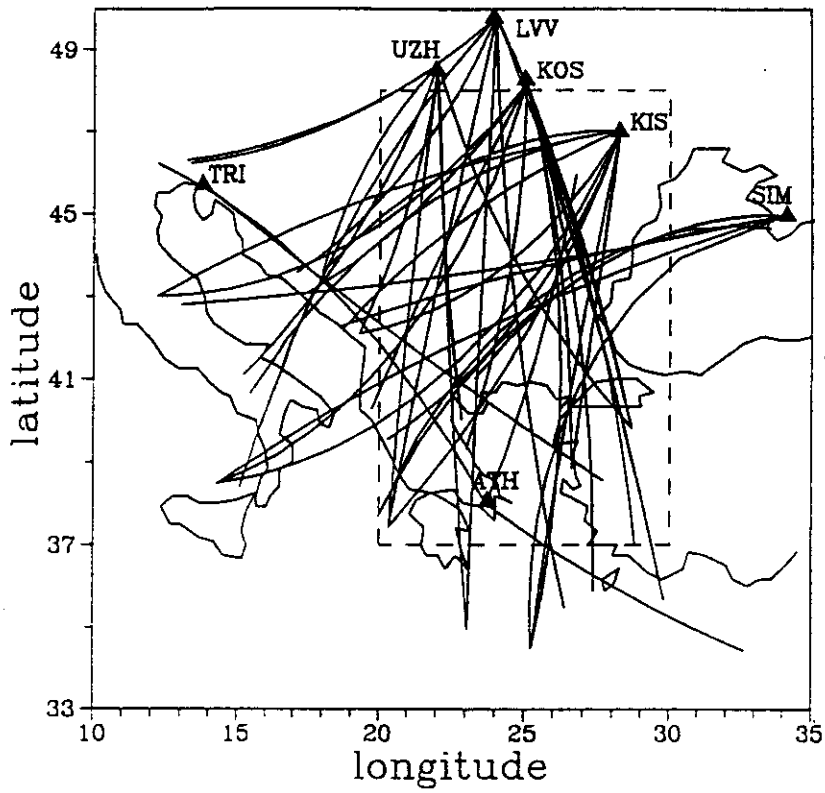

Figure 8

Pattern of Rayleigh wave paths for which azimuth anomalies have been determined from the polarization analysis by NESTEROV and YANOVSKAYA (1988). The paths are shown schematically as circular arcs with the observed azimuthal anomalies at the stations. The dashed line contours the area where the solution has been calculated.

The solution is affected more strongly by regularization parameter γ (see eq. (21)), which depends on the standard error of the data. Since this error was not known, and it is presumed to be large enough, parameter γ was chosen arbitrarily, such that the lateral velocity variations in the solution would be realistic, i.e., not exceedingly large, which could occur if the regularization parameter were taken too small. Therefore, the solution should be regarded only as an approximation to the real velocity distribution, and only locations of low and high velocity zones may be obtained with confidence from these data.

Since the regularization parameter was assumed to be large enough, a strong reduction of the data variance could not be expected. Indeed, the square root of the variance was reduced from 15.4° to 8.3° for $T = 10$ s and from 13.1° to 7.7° for $T = 25$ s.

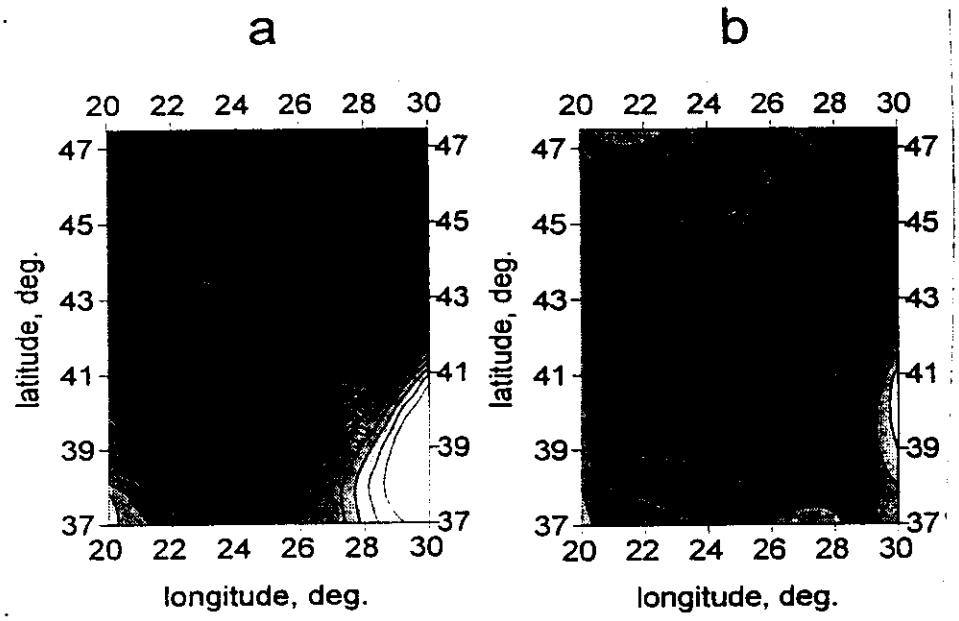

Figure 9
Lateral phase velocity distribution obtained from azimuth anomalies (a) and group-velocity distribution obtained from group-velocity data (b) for period $T = 10$ s.

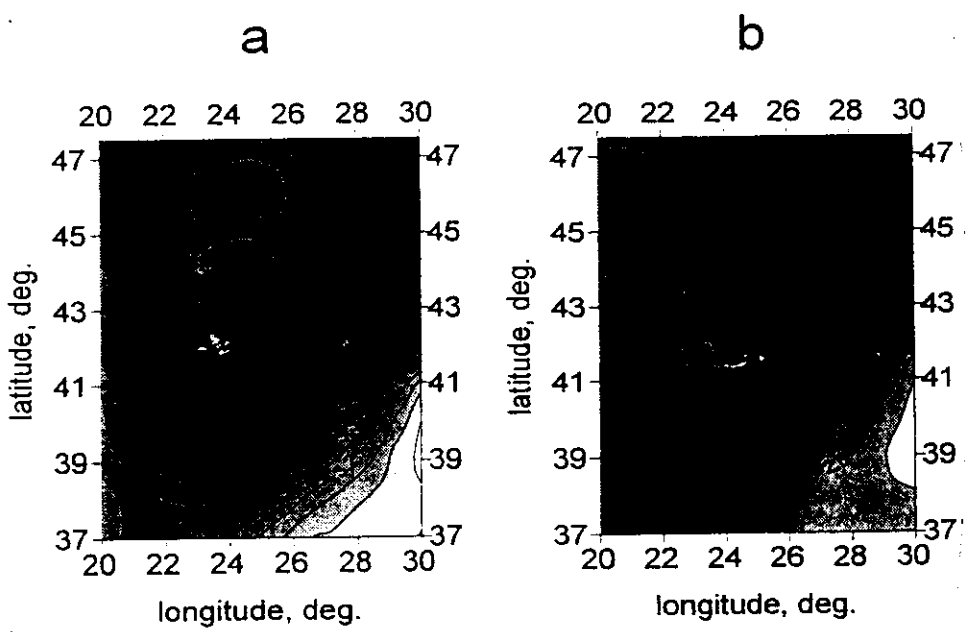

Figure 10
The same as Figure 9 for period $T = 25$ s.

The phase velocity distributions are shown in Figures 9a and 10a. Unfortunately, we have no data pertaining to the phase velocities in the region, therefore the results could not be compared with these quantities. The only data available for the area are group-velocity distributions obtained from travel-time tomography. Certainly, a strict correlation between group and phase velocities should not be expected, though surface wave studies reveal a correlation between group velocities and the velocity structure similar to that for phase velocities. In particular, in areas with a thick sedimentary layer we observe low group velocities (and also low phase velocities), and if the sedimentary layer vanishes, both group and phase velocities are high.

Group-velocity distributions for the same periods within the same area are delineated in Figures 9b and 10b (NESTEROV and YANOVSKAYA, 1991; YANOVSKAYA, 1993). A striking similarity between the phase and group velocities is observed: the low and high velocity zones are located in the same places, with even their shapes similar. However, to capably draw conclusions on the reality of the phase velocity solution from azimuthal data, it is necessary to compare it with the phase—rather than the group-velocity distribution.

Conclusions

An approximate relationship between the lateral velocity variations and azimuth anomaly has been obtained for the case of constant velocity in the starting model. In spite of noticeable errors due to the approximation, this relationship may be used for solving tomography problems, because the errors do not display any systematic trend and, consequently, may be included among observation errors. This is confirmed by numerical examples.

It is evident that more reliable and detailed results may be obtained by the joint inversion of azimuth anomalies and the data on phase velocities, since they contain different information on the phase-velocity distribution.

The main problem in azimuthal tomography is the determination of azimuthal anomalies from observations: the superposition of Rayleigh and Love waves, as well as multipathing, lead to large errors in azimuths.

Acknowledgements

This research was supported by International Scientific Foundation Grants NWV000/NWV300 and Russian Foundation of Fundamental Research Grant 94-05-17422a.

REFERENCES

FARRA, V., and LE BÉGAT, S. (1995), *Sensitivity of qP-wave Travel Times and Polarization Vectors to Heterogeneity, Anisotropy and Interfaces*, Geophi J. Int. *121* (2), 371–384.

HU, G., and MENKE, M. (1992), *Formal Inversion of Laterally Heterogeneous Velocity Structure from P-wave Polarization Data*, Geophys. J. Int. *110*, 63–69.

HU, G., MENKE, W., and POWELL, C. (1994), *Polarization Tomography for P-wave Velocity Structure in Southern California*, JGR *99* (B8), 15,245–15,256.

JULIAN, B. R., and GUBBINS, D. (1977), *Three-dimensional Seismic Ray Tracing*, J. Geophys. *43*, 95–113.

LANDER, A. V. (1984), *Surface Wave Anomalous Occurrences in Northeast Eurasia and their Relation to the Momsky Rift Area*, Computational Seismology, *16*, 127–155 (in Russian).

LASKE, G., MASTERS, G., and ZURN, W. (1994), *Frequency-dependent Polarization Measurements of Long-period Surface Waves and their Implications for Global Phase-velocity Maps*, Phys. Earth and Planet. Inter. *84*, 111–137.

LERNER-LAM, A. L., and PARK, J. J. (1989), *Frequency-dependent Refraction and Multipathing of 10–100 Second Surface Waves in the Western Pacific*, Geophys. Res. Lett. *16*, 527–530.

LEVSHIN, A. L., RITZWOLLER, M. H., and RATNIKOVA, L. I. (1994), *The Nature and Cause of Polarization Anomalies of Surface Waves Crossing Northern and Central Eurasia*, Geophys. J. Int. *117*, 577–590.

NESTEROV, A. N., and YANOVSKAYA, T. B. (1988), *Lateral Lithosphere Inhomogeneities in Southeastern Europe from Surface Wave Observations*, Izv. AN SSSR, Fizika Zemli *11*, 3–15 (in Russian).

NESTEROV, A. N., and YANOVSKAYA, T. B. (1991), *Inferences on Lithospheric Structure in Southeastern Europe from Surface Wave Observations*, XXII General Assembly ESC, Proceedings and Activity Report 1988–1990, *1*, 93–98.

SNIEDER, R., and SPENCER, C. (1993), *A Unified Approach to Ray Bending, Ray Perturbation and Paraxial Ray Theories*, Geophys. J. Int. *115*, 456–470.

TARANTOLA, A., and NERSESSIAN, A. (1984), *Three-dimensional Inversion Without Blocks*, Geophys. J. Roy. Astr. Soc. *76*, 299–306.

YANOVSKAYA, T. B., and DITMAR, P. G. (1990), *Smoothness Criteria in Surface Wave Tomography*, Geophys. J. Int. *102*, 63–72.

YANOVSKAYA, T. B. (1993), *Study of azimuthal anisotropy and lateral heterogeneities of seismic zones*. In *Mathematical Modelling of Seismotectonic Processes in Lithosphere Aimed to the Problem of Earthquake Prediction*, *1*, 93–105 (in Russian).

(Received July 17, 1995, revised December 21, 1995, accepted January 22, 1996)