Inverse Problems of Seismology (Structural Review)

V. I. Keilis-Borok and T. B. Yanovskaja

Summary

The inverse problem of seismology is formulated as the determination of the set of cross-sections fitting all given observations.

The schemes of solution of this problem by the trial-and-error method and the inversion method are described. Mathematical and computational problems involved in the solution are outlined. Not all these problems are solved at the present time. Among them optimal parametrization of cross-section and the determination of a probability distribution function for cross-section are the most urgent problems. But with certain limitations the described scheme of a trial-and-error method for joint interpretation of surface and body waves is in routine operation now. Results obtained are described in papers [2, 4, 6, 15, 22, 23]. Usually it is the non-uniqueness that strikes us when we see the solution of the inverse problem. But there is another side to the matter: if we do not look for an excessively detailed cross-section we may obtain quite definite conclusions even from very rough observations which are commonly regarded as useless.

1. Introduction

The inverse problem of seismology is formulated as follows:

Given: (a) some combination of observations on body and surface seismic waves and on free oscillation of the Earth; (b) postulated data on Earth constitution, that is limitations regarding the cross-section as well as the relationship between various parameters.

To be found: a set of cross-sections which are in agreement with the given data allowing for their precision.

The above formulation of the problem is evident but not quite usual. In common practice, a single cross-section is being looked for and different single cross-sections from a whole set may be obtained in different investigations by chance especially if different waves have been used.

There are two approaches to the solution of the inverse problem of seismology:

(a) the inversion of records, or of parts of records, as a function of time, without separation of single waves.

(b) the inversion of some functions characterizing the laws of propagation of single waves.

The following functions can be considered: hodographs (or travel-time curves) $t(\Delta, h)$; apparent velocities $dt(\Delta, h)/d\Delta$; amplitude–distance curves $A(\Delta, h, T)$; phase and group velocities $c_p(T)$ and $C_p(T)$; periods $T_k$ and decrement of attenuation $\lambda_k$ of the free oscillations of the Earth.
Here \( t \) is travel-time, \( \Delta \) is epicentral distance, \( h \) is depth of focus, \( A \) is amplitude, \( T \) is period of oscillations, and \( k \) is number of mode.

Two methods are possible in each approach:

(a) Trial-and-error method to find the cross-sections for which the theoretically calculated properties of waves are in agreement with observations.

(b) Direct inversion of data; that is, a solution of problem inverse in terms of mathematics.

The first approach, the inversion of records, has been developed only in a few recent papers [18, 19, 20]. The technique of solution of the one-dimensional inverse problem is well developed in seismic prospecting, but the one-dimensional problem is seemingly of little interest to seismology. This technique is formally generalized to a three-dimensional case when the medium is horizontally homogeneous [5]. But to use this generalization one needs unreasonably exact and detailed observations. Only the second approach, i.e. inversion of characteristics of propagation of separate waves is considered in this paper.

2. Trial-and-error method

The general scheme of the method is given in Fig. 1(a). Solid lines indicate the repetitive cycles, and dotted lines indicate the once-round ones.

*Summary of observations* needs no comment.

The raw observations are to be presented here, i.e. the clouds of measured dots, without drawing the mentioned functions and without identification of waves; they have to be divided only into body, Rayleigh and Love waves and into spheroidal and torsional free oscillations. Only the simplest corrections, which make the dots comparable, are allowed here.

![Diagram](https://example.com/diagram.png)

**Fig. 1.** The general scheme of the trial-and-error method (a) and the direct inversion method (b).
For example, the arrival times are given as a cloud of dots in the \((t, \Delta)\) plane for each \(h\), without drawing the hodograph \(t(\Delta, h)\), even less its loops; velocities \(c_k\) or \(C_k\) are given as a cloud of dots in the \((c, T)\) or \((C, T)\) plane, without drawing the dispersion curve or indicating the mode number \(k\). The error should be estimated for each dot.

**Summary of a priori data**, i.e. of the limitations on cross-section, which follow from previous investigations of the same cross-section, and from physical considerations. This summary should include also the possible relation between different properties—elasticity, density, attenuation, etc. [14].

**Parametrization of cross-section** means the representation of cross-section through a finite number of numerical parameters, and the indication of a priori limits for them. This operation is based on a priori data and on the preliminary qualitative analysis of given observations. One example of such an analysis is given below in Figs. 3 and 4. The pattern of the cloud of observational dots indicates how detailed can be the representation of the cross-section.

Optimal parameters are those which can be most precisely determined from the given observations. The problem of the choice of optimal parameters is not solved yet. To illustrate the formulation of this problem, we shall describe some preliminary results of [8,9] on determination of the cross-section \(v(r)\) from a given hodograph \(t(\Delta, h)\) for a spherically symmetrical model of the Earth \((r < R)\) in the presence of waveguides. Here \(v\) is the wave-velocity, \(r\) the distance from the centre of the Earth. The transformation

\[
x = \frac{R}{v(R)} \Delta, \quad z = \frac{R}{v(R)} \ln \frac{R}{r}, \quad a(z) = \frac{v}{v(R)} \exp \left\{ - \frac{v(R)}{R} z \right\}
\]

reduces this problem to determination of velocity \(a(z)\) in the half-plane \(z \geq 0, -\infty < x < \infty\). It is shown by [8,9] that \(a(z)\) and, consequently, \(v(r)\) are non-uniquely determined even by an exact hodograph in the presence of waveguides. In that case the hodograph determines not the unique cross-section \(a(z)\) but some region in the \((a, z)\) plane inside which the cross-section lies. One example of such a region is shown on Fig. 2. This region corresponds to Gutenberg’s surface-focus hodograph [10], assumed as being exact.

However, some characteristics of the cross-section are determined uniquely by the hodograph, for example the following.

If we have only a travel-time curve for surface focus \((h = 0)\) with some shadow zones, then the following integrals depending on \(a(z)\) within the waveguides must be constant:

\[
\int_{Z_i}^{\bar{Z}_i} \sqrt{(a^{-2}(z) - p_i^2)} dz = \sigma_i.
\]

Here \(Z_i\) and \(\bar{Z}_i\) are the depths of upper and lower boundaries of the \(i\)th waveguide, \(p_i = a^{-1}(Z_i) = a^{-1}(\bar{Z}_i)\), and the values of \(\sigma_i\) are determined from the given travel-time curve in the following way: if we draw the tangents to the ends of the hodograph before and after the \(i\)th shadow zone, \(\sigma_i\) will be the intercept between these parallel tangents on the \(t\)-axis.

If in addition we have a set of hodographs for focal depths \(\bar{Z}_i \leq h_i < Z_{i+1}\) the following additional characteristics are determined:

\[
H_i(\bar{a}) = \text{mes} \{ z, Z_i \leq z \leq \bar{Z}_i, a(z) \leq \bar{a} \}
\]

(3)
Fig. 2. The region of possible $P$-velocity cross-sections below the crust, corresponding to Gutenberg's hodograph (10) (after V. M. Markushevich). Vertical lines extend to $V = 0$.

i.e. $H_i(\bar{a})$ is the total length of interval inside the $i$th waveguide, where the velocity is less than $\bar{a}$. More details can be found in [8]. Precisely similar hodographs correspond to different cross-sections, for which $H_i(\bar{a})$ and $\sigma_i$ are equal. Evidently, parameters of cross-section should be chosen in such a way that their different combinations would correspond to different $H_i(\bar{a})$, $\sigma_i$.

The further choice of optimal parameters can be illustrated by the following results from [9] on the investigation of cross-section above the focus. Let us consider the parametrical form of the hodograph

$$t = t(p), \quad x = x(p), \quad p = \frac{dt}{dx},$$

and the moments of the hodograph

$$\beta_i = \int_0^1 \theta^i t(\theta P) d\theta,$$  (5)

$P$ being the value of $p$ at the inflection point. If several of the first $\beta_i$, say $\beta_0$, $\beta_1$, ..., $\beta_n$ are known, the same number of values $b_i$,

$$b_i = \int_1^2 \phi^i dH_\tau(\phi), \quad dH_\tau = \frac{P\phi}{4\sqrt{(\phi - 1)}} dH_b \left(\frac{2\sqrt{(\phi - 1)}}{P\phi}\right)$$  (6)
can be determined using the following equations [9]:

$$
\beta_{2k+1} = \frac{k!}{(2k+1)!} \sum_{i=1}^{k+1} \frac{(2k-i+1)!}{(k-i+1)!} b_i.
$$

Values of $\beta_i$ are some characteristics of the cross-section—simultaneously with $b_i$. For a given error of arrival times the error of $\beta_i$ increases with $i$. That is why only a limited number of $\beta_i$ and consequently, $b_i$, can be determined with a given accuracy. This number increases with accuracy of observations; at the limit, the infinite set of $b_i$ determines the cross-section.

It is evident that $b_i$ are the optimal parameters of cross-section which should be determined from such observations. Their physical significance is not clear. But a set of cross-sections corresponding to each combination of $b_i$ can be calculated afterwards for simplicity.

However, the solution of (7) involves some computation difficulties; other parameters may be optimal for other kinds of observations. That is why the problem of the choice of optimal parameters is not solved yet. One possible method of its solution will be mentioned a little later, in connection with the last operation.

In routine practice the cross-section is usually divided into layers with constant velocity or velocity-gradient. Some of the following parameters are used: depth of the boundaries of the layers; velocity, velocity gradient, jump of velocity or jump of velocity gradient on the boundaries; average velocity on some depth-intervals. To ensure that different combinations of such parameters will correspond to different hodographs, one should assume a standard form of $a(z)$ in each waveguide. A set of other possible forms can be calculated afterwards, using (2) and (3). In the results of this operation we represent the cross-section as a point in the space of its unknown parameters. Their limits determine the region where this point lies. Our problem is to narrow this region as far as the observations allow. It is being done by the three following operations.

Points of the region mentioned above (i.e. possible cross-sections) are tried in succession. The data known from observations are computed theoretically for each point. The discrepancy between theoretical and observed data is estimated.

The solution of our inverse problem is a region where this discrepancy is sufficiently small. Our problem is reduced, therefore, to the determination of the region of the minimum of some function (the discrepancy mentioned) in a multidimensional space of unknown parameters of cross-section.

The flow of cross-sections is organized in accordance with the accepted method of search for a minimum. Three methods are widely employed: the Monte-Carlo method, the descent along the gradient and the random search method. They are compared in [4].

The descent along the gradient happens to be effective only for a small number of unknown parameters.

The method of random search is effective far from the minimum; but the search gives way to casual wandering in the vicinity of the minimum.

Strange as it may seem, the Monte-Carlo method proves to be preferable; however, it also possesses a great disadvantage for the results of the trials already made are not used in the next trial. On gaining the minimum we leave it at once. That is why only a small percentage of trials prove to be successful.

The ‘hedgehog method’ developed by V. Valus seems to be the most promising one. It consists of the following. A single point of the minimum region

$$y(y_1, y_2, \ldots, y_n),$$
\( y_i \) being the parameters of the cross-section, is found by the Monte-Carlo or some other technique. Then the neighbouring points are tried,

\[
y + \sum_i x_i d y_i,
\]

where \( x_i = 0 \) or 1 and certain combinations of \( x_i \) are picked by turns.

Points which fall within this minimum region are selected. The same procedure is applied to every selected point until the whole region is covered.

Upon this, we return to the Monte-Carlo technique (omitting, of course, the found region from further investigation) and look for another minimum region, and so on.

Direct problem computation—the theoretical computation of the data which are known from observations—needs no comment. Practically, it can be done for any kind of data, but only if the model of the cross-section which we try is horizontally homogeneous.

The programs used in our works are described in [2, 17, 24].

Comparison of computations and observations. The choice of the measure of discrepancy between them is the most important but inadequately developed problem. This measure should answer the purpose of investigation: it must depend heavily on those properties of cross-section which we wish to investigate, and depend little on other properties. Unfortunately we do not always know in advance what we are investigating the cross-section for.

Ideally, the measure should be the probability that the discrepancy between observations and computations is random. This probability, after proper normalization, could be considered as the probability of the corresponding cross-section. Then the solution of our inverse problem could be represented in a most natural form—as a statistical density of distribution of the unknown cross-section (more exactly, of the corresponding vector in the space of unknown parameters). Afterwards a confidence region could be determined. The method of determination of this distribution has not yet been developed. In routine practice we followed the usual method: comparison of theoretical curves with corresponding curves obtained by smoothing of observations. The curves were represented by a set of ordinates. The measure of discrepancy was chosen by numerical experiments as some combination of the usual measures; standard and maximal deviations, correlation coefficient, etc.; one such experiment is described in [4].

This practice is far from being optimal, but it gives comparatively good results.

Compact description of the found cross-sections means the determination of their common geometrical properties. A trivial example is the average velocity; it is usually approximately the same for all cross-sections [2, 23].

If the parametrization were optimal, in the sense indicated on page 225, we would automatically obtain the compact description; it would be given by the intervals of values of some parameters for all cross-sections found. And vice versa, the common properties of the cross-sections are the optimal parameters which should be introduced from the beginning; so, after this operation, the parametrization of cross-section can be improved.

The compact description of cross-sections can be found using the methods of statistics or recognition. But at the present time there is no experience in application of these methods to our problem. In routine practice we described the solution by possible intervals for average velocity, for velocities at some depths and for depth of discontinuities.

The described trial-and-error method with forced simplifications, mentioned above, is in routine operation now for joint interpretation of hodographs and amplitudes of body waves, dispersion of surface waves and free oscillations. In order to add any other data—gravitational, thermal, etc.—it is necessary only to include in

Downloaded from https://academic.oup.com/gji/article-abstract/13/1-3/223/919956 by guest on 22 January 2019
the operation 'direct problem computation', the subroutine for computing the corresponding data theoretically, for a given cross-section, and to formulate the algorithm for comparison.

3. Inversion method

The general scheme of this method is given on Fig. 1(b).

Four operations are the same as in the trial-and-error method, and we have to describe only another three.

Parametrization of curves means the representation of each curve by a finite number of parameters and the indication of a priori limits for them. This operation is based on the general theory of these curves, on a priori data and on the pattern of observational dots.

The general theoretical properties of hodographs are given in [8, 9]. For dispersion curves they are not yet formulated.

The problem of optimal parametrization of curves is equivalent to that of cross-sections and is not yet solved. In routine practice we represent each curve by some set of ordinates or by the parameters of straight lines and parabolae which approximate separate pieces of the curve.

The flow of curves means here the consecutive picking of various combinations of corresponding parameters. All possible variants of wave identification should be considered, of course, for each variant of the curve.

One example is shown in Figs. 3 and 4.

Fig. 3 shows the summary of observations on waves propagated through the Earth's core. Dots are the data from [1, 11]. Lines show the hodograph of Jeffreys and Bullen [12]; the corresponding dots from which the lines were constructed are now not available.

![Figure 3](https://academic.oup.com/gji/article-abstract/13/1-3/223/919956/fig3.jpg)

Fig. 3. Observations on core waves: lines, after (12); dots, after (1); circles, after (11).
FIG. 4. Some types of hodograph, fitting qualitatively the data shown in Fig. 3. Dotted lines are the hodographs of reflected waves. The corresponding types of cross-sections are shown below:

(a), (b)—two-layered core. (c), (d), (e), (f)—three-layered core.

The sequence of parts of the hodograph corresponding to decrease of the angle to the vertical of the seismic ray at the focus is indicated by the sequence of letters.

Fig. 4 shows different types of hodographs which could fit in with the given data; the corresponding types of cross-sections are drawn below. More detailed representation of the core is hardly possible with the given data.

To get the complete solution of our problem we should investigate also some other types of hodographs and for each type draw a set of variants. Fig. 5(a) shows in greater detail one of the variants of the type similar to that shown in Fig. 4(c) but without sharp discontinuities and the corresponding cross-section, obtained by the Herglotz-Wiechert method. The theoretical amplitude-distance curve for this cross-section is shown in Fig. 5(b); it is computed neglecting the attenuation.

The comparison of a given variant of the curve with observations is the same operation as in the trial-and-error method. The only difference is that now we compare with the observations not the curve computed for a given cross-section, but the curve drawn directly in the cloud of the observed dots.

Inversion of the curve into a cross-section (or parameters of the curve into parameters of the cross-section) is the next operation. It is applied to all variants of curve which fit in with the observations.

The method of inversion is developed only for hodographs [8].
The last operation—compact description of cross-section—has already been mentioned. For improvement of parametrization of curves we have to include in this operation the compact description of curves which are in agreement with observations.

The inversion method described here is practically undeveloped, but the pioneering paper is published [13]. This paper gives the method of drawing the variant of a hodograph with the maximum statistical likelihood ratio. Such variant may have no practical advantage compared with other variants, which are of a slightly lesser likelihood. The determination of a distribution function and confidence region for a curve in clouds of dots remains the key problem in this method also, as well as the optimal parametrization.

In practice, inasmuch as this problem is not solved, the simplest procedure seems to be satisfactory: to approximate hodographs by a finite number of straight lines and second order curves and to pick in turn variants of the hodograph with small standard and maximal deviations from observed travel times and apparent velocities.
Fig. 6. Theoretical amplitude-distance curves for Jeffreys' velocity distribution (surface focus) and attenuation given in [3]. Curves with no indication of period $T$ correspond to the model without attenuation. Primes correspond to reflections from the $M$-discontinuity. After T. B. Yanovskaja and L. M. Kininchuk.
4. Comparison of both methods described

The advantage of the trial-and-error method is the possibility of analyzing jointly the largest amount of different data—any for which the direct problem can be calculated theoretically. One example is the joint interpretation of the first arrival delay and of Bouguer anomaly [21]. Another example is the joint analysis of seismic and temperature data [16]. The temperature distributions in the crust were calculated in [16] for different cross-sections of the crust, obtained in [15].

Differences in temperature amount to 100° at a depth of 20 km, and to 500° at the bottom of the crust. Some cross-sections give a strong negative temperature gradient and possibly may be rejected.

The main advantage of the inversion method is the smaller amount of computation time necessary.

At present, when it is developed only for travel-times of body waves, the combined method may be useful: to get a set of cross-sections by inversion of travel-times and then, by trial-and-error technique, to select cross-sections which fit the other observations as well.

5. The qualitative approach to the inverse problem

This approach is unduly neglected in seismology. An example of its possibilities is a qualitative survey of the body wave pattern as given in Jeffreys–Bullen table [12].

Fig. 6 shows theoretical amplitude–distance curves for some body waves mentioned in these tables. They are calculated for Jeffreys’ model of the Earth, and a surface source. For the normal source the picture is practically the same. One ought to pay attention to the following:

(a) Waves reflected from the Mohorovičić discontinuity are seemingly observed instead of SS, SSS in zones of minima of respective curves (the minimum is due to reflection).

(b) PPP and SSS waves could not be observed, if for Q we accept the data derived in [3] from surface waves.

(c) The theoretical curve for P-waves has none of the fluctuations of the empirical curve of Gutenberg. If these fluctuations are reliable, they indicate fluctuations of the velocity gradient.

Other examples of the qualitative approach can be found in [2, 7]: the form of the envelope of surface waves records gives an indication of the presence or absence of a waveguide.

Soviet Geophysical Committee,

Molodezhnaya 3,

Moscow B–296,

U.S.S.R.

References


